

Carter's energy message

President Carter has emerged from his self-imposed period of isolation to announce a six-point plan for cutting dependence on foreign oil imports. Those who expected his long silence to be followed by some dramatic gesture were to be disappointed; the content of his speech to the nation was confined to proposals which had received wide currency in recent months — holding oil imports to 1977 levels (themselves the highest ever) and paring down imports by a half over the next decade, developing synthetic fuels, making a more positive start on conservation and so on. Nothing was said about nuclear power. No imaginative political moves were made — but this is not President Carter's style.

At least Carter's aims are relatively modest. At the time of the 1973 energy crisis President Nixon wheeled out a plan, 'Project Independence', of which was intended to make the United States independent of foreign energy suppliers by the year 1980.

Perhaps the most eye-catching proposal is that alter-

native fuels, notably synthetic fuels, should account for 2.5 million barrels per day by the year 1990. This would be rather more than a tenth of present-day oil consumption. Even this (as David Dickson points out on page 181) is little more than catching on to a recent optimistic public mood that expensive synthetic fuels will eventually be acceptable in the market place. And environmental objections, including those concerning carbon dioxide (see also page 189), will have to take a back seat.

President Carter has chosen to couple his energy message with a deeper warning to the American people that there is a crisis of confidence, moral and spiritual, in the nation. There may be a crisis of confidence in the presidency, and certainly Americans are worried about their future industrial competitiveness, but to bring in moral and spiritual values is surely to confuse the real message: that a whole raft of institutional and social changes are going to be necessary to meet a new era of energy austerity. □

Aldermaston goes a little public

The invitation for a number of members of the press to visit Aldermaston, the centre of Britain's atomic weapons' research, was an unexpected and pleasant surprise. Defence research establishments rarely open their doors, and as far as anyone could remember this was the first time a contingent from the press had been inside the security fence. The purpose of the visit was, from Aldermaston's point of view, to demonstrate their new laser compression facility recently opened by the Queen — a facility which seemed strangely similar to another laser compression facility just 30 kilometres away at the Rutherford Laboratory. Why can't the weapons people and the academics pool their resources (the price tag is, after all, several million pounds) and share a bigger and better facility? The Americans seem to have far less inhibitions about doing defence and academic work in the same building.

But the purpose of the visit, from many of the journalists' point of view, was not so much to look at a laser, nor even at the splendid display of non-weapons-related research gathered in a giant marquee and discretely draped with reminders that a pay dispute was still on. Rather, it was a unique opportunity to fire at the Director of the establishment a whole range of questions about broader issues concerned with nuclear weapons.

The most persistent questioning concerned the return to action of the plutonium facility which had to be closed last August when some workers in it were found to have been exposed to higher than permitted doses of radiation. It seems that as yet a substantial fraction of the facility is still not fully in operation. One of the big problems is manpower — not just skilled craftsmen but also health physicists prescribed by the Pochin report.

Even on this matter the Director and his colleagues were the very soul of discretion. When it came to questions such as 'What is happening under the Polaris improvement programme?' 'What is going to happen to the US/UK agreement on nuclear exchanges, due for renewal this year' and 'How would a comprehensive test ban pose problems for nuclear weapons stockpiles', the discretion was absolute — each was answered with a 'no comment'.

It would have been a somewhat disgruntled team of journalists which would have left Aldermaston under these circumstances were it not that the message did seem to have registered that the lack of public information on the British nuclear deterrent, particularly at a time at which major decisions are being made, is a matter of some concern. Promises were made that the Ministry of Defence would look into giving the press a fuller briefing in the near future on nuclear policy. The sooner the better. □



Sparks continue to fly in low-level radiation row

David Dickson describes a conflict in paradigms behind differing interpretations of the health hazards of nuclear power

Two and a half months after the National Academy of Sciences submitted to the Environmental Protection Agency the "final draft" of a long-awaited report on the biological effects of ionising radiation, an *ad hoc* committee within the academy is still meeting to re-write the most controversial section of the report, that discussing the appropriate model for interpreting the dose-response relationship for low levels of human exposure.

With public concern at the possible health hazards of nuclear power fuelled by the recent accident at Three Mile Island — and the nuclear industry no less concerned about meeting the costs of a possible tightening of occupational exposure levels — the issue is inevitably a hotly-debated one. Hopes that the temperature might be reduced by referring the issue to the academy, however, have been frustrated by a dispute over the interpretation of scientific evidence as heated as the public controversy outside.

The scientific dispute emerged briefly into public view at a press conference in May, when presentation of the draft conclusions of the academy study was accompanied by a strongly-worded minority statement claiming that the report's conclusions could lead to "detrimental apprehension" over radiation hazards; when additional members of the drafting committee indicated their own reservations, academy president Dr Philip Handler decided that the disputed section should be re-written.

Controversy already surrounds two other academy reports on politically sensitive issues, one on methods of nuclear waste management which was withheld from publication partly at the request of the Department of Energy, the other a massive study on energy policy which is already two years behind schedule. And Dr Handler has now announced that top NAS officials will be taking a close look at procedures for handling such issues over the summer "to see that these type of problems don't happen again."

Concern over the health effects of radiation is as old as knowledge of radiation itself. Until the mid-1960s, the main issue of controversy was whether a threshold exists below which radiation can be assumed to have negligible effect; since then, most scientists have accepted that there is probably no such threshold, and the debate has been over the shape of the dose-response relationship at low levels of exposure.

In a report in 1972, the academy's Committee on the Biological Effects of Radiation (BEIR) suggested that a linear hypothesis, extrapolating backwards from the known relationship at high levels, was

the most appropriate to adopt for regulatory purposes — given all the uncertainties involved — and this has been accepted by various standard setting bodies.

However, this hypothesis has been attacked from both sides. Supporters of the nuclear industry claim that it is too restrictive, considerably overstating the risks; on the other hand studies of nuclear workers and others have been reported which, industry critics claim, show even existing exposure levels to be too high.

The BEIR committee, in its new report, sticks to the linear hypothesis as the most "prudent" assumption. Within the committee, however, a fierce battle has raged over whether, and if so to what extent, this assumption overstates the "true" risk. Both sides in the dispute agree that there is insufficient human data to provide statistically valid risk estimates for exposures of less than about 10 rads; where they disagree, with almost religious fervour, is what type of models can be used to fill the gap.

Supporters of the linear hypothesis tend to rely predominantly on epidemiological and statistical arguments. They claim, firstly that there is sufficient evidence from other studies, such as the effects of the medical uses of radiation, to make the linear hypothesis plausible; and secondly that the lack of human data makes it statistically impossible to reach any more precise assessment.

Their opponents, in contrast, tend to rely more heavily on radiobiological data obtained from studies of the effects of radiation on lower organisms. In the minority report attached to the draft of the committee's report, it is stated that "even if the linear model were in fact to afford a substantially better fit to epidemiological data, this would be insufficient justification for rejection of models that are consistent with a vast body of radiobiological evidence."

According to the latter arguments, a more appropriate dose-response model (at least for "low energy transfer" radiation, such as gamma rays) would be a quadratic one, possibly with a linear component. The linear hypothesis, they suggest, should only be accepted as an upper limit, with suitable recognition of a low limit of equal validity. "It should then be stated that the true risk is between these limits, and very likely near the lower limit".

Radiobiological data is also given greater prominence in deliberations within the National Council for Radiobiological Protection, a body which, although private, has been attacked by critics for its links to the nuclear establishment. Dr Victor Bond, for example, chairman of an NCRP committee which is soon to

produce a report on low level radiation expected to reach different conclusions to the BEIR committee, told a Congressional committee last month that there were "good biological reasons" to support the lower estimates.

Dr Bond referred in particular to studies such as those carried by Dr Harald Rossi, one of the most outspoken critics of the linear hypothesis on the BEIR committee, and colleagues at the Columbia University College of Physicians and Surgeons in New York, on genetic changes in cells of the plant *Tradescantia*. "Radiobiology in simple systems is important to the establishment of the most-probable dose-effect relationships for the human being" he said, adding that the linear hypothesis for low level radiation was "biologically unsubstantiated".

Nowhere is the difference between the supporters and the critics of the linear hypothesis more marked, however, than in their different interpretations of studies on survivors of the atomic bombs dropped on Hiroshima and Nagasaki.

The critics point for support of their curvilinear hypothesis to studies of the mortality rates from all forms of cancer carried out between 1950 and 1974 on 80,000 survivors in the two cities. Concentrating in particular on the Nagasaki mortality data, this reveals an effect "about ten times lower than that based on proportional projections from persons receiving high X-ray doses" according to Dr Edward Webster of Harvard Medical School, co-author with Dr Rossi of the original minority report.

Supporters of the linear hypothesis, however, look at the Japanese data differently. They criticise the mortality data for its lack of statistical precision (in particular in the light of uncertainties over the relative effects of gamma and neutron components), arguing that "it is difficult to define a functional form in the dose response relationship" and that a linear-quadratic curve "is not demonstrably superior to a simple linear fit."

As an alternative, the original draft of the report suggests that it is more useful to look at data on cancer incidence rather than on mortality; and that a registry of solid tumours among Nagasaki survivors — evidence whose validity is questioned by the critics — supports their general hypothesis.

A further dimension to the dispute — and, one that some argue, is largely fuelling it — has been a sharp difference in attitude towards the nuclear power industry. Dr Edward Radford of Pittsburgh University, chairman of the BEIR committee, has made little secret of his personal views that current permitted exposure levels for nuclear workers are unnecessarily and perhaps dangerously high, suggesting that exposure levels be

reduced by a factor of up to ten.

"The purpose of requiring more restrictive occupational exposure limits is to provide clear design criteria to allow engineers to design facilities to meet these kinds of standards, rather than rely on the somewhat nebulous and legally dangerous to the industry 'as low as reasonably achievable' concept," he says.

The dissenting group on the BEIR committee has made it equally clear that it feels too restrictive exposure levels could place an unnecessarily heavy burden on the nuclear industry. The original minority report warns that overestimates of radiation hazard could result in "serious detriment" to national energy policy. "If the guide line levels were reduced in the way

[Radford] wants them, there would not be any nuclear industry at all" one committee member has been quoted as saying.

The BEIR committee dispute has been characterised as a conventional scientific disagreement. "Such controversies have existed within the scientific community since time immemorial. We will truly be in trouble when scientists passively accept each other's work without challenge," says Dr Sidney Marks of the Battelle Memorial Institute's Pacific North-West Laboratory.

In this case, however, the incentive has been primarily provided not by a desire for knowledge, but by a specific regulatory need; and the resulting impasse therefore carries both scientific and political

dimensions. Initially the academy had hoped that this could be accommodated in the minority report; however when it emerged that over half of the BEIR committee's somatic effects subcommittee shared the reservations about the degree of support given to the linear hypothesis, it was decided that a redrafting of the whole section to produce "a more balanced expression of differing interpretations" would be more appropriate.

The new version will, it is hoped, be ready for re-submission to the full committee within a few weeks. The academy still says that "dissenting views will be published in the final report"; so whatever is agreed is unlikely to be taken as the last word. □

US scientists warn of environmental dangers from synthetic fuels

As President Carter was preparing the details of a major national programme to stimulate the production of synthetic fuels, announced in his energy speech last Sunday, the Council of Environmental Quality released a report warning that a rapid increase in the use of synthetic fuels could have serious environmental consequences through speeding up the accumulation of CO₂ in the atmosphere (see also page 189).

Synthetic fuels, previously a relatively unfashionable area of research, have suddenly been receiving the enthusiastic attention both of the President and of the US Congress.

Despite wide uncertainties over both economic and technical viability, the House of Representatives has already passed a bill authorising a crash development programme aimed at a target of producing 2 million barrels of synthetic fuel a day by the year 1990. And in his energy speech, delivered after almost two weeks of closeted discussion with his top advisers, President Carter proposed a programme that would develop America's alternative fuel sources from coal, oil-shale, plant products and the Sun to replace 2.5 million barrels a day of imported oil by 1990.

However both environmentalists and scientists have expressed concern at the environmental implications of such programmes, urging the President to concentrate on means for achieving energy conservation rather than the search for substitute fuels.

In a report to the CEQ, the four scientists say that the production of CO₂ resulting from the use of synthetic fuels (such as gas or oil) made from coal is almost 50% higher than the CO₂ emissions resulting from the production of energy directly from coal, and twice as high as the result of obtaining an equivalent amount of energy from natural gas.

"It is our conviction that an appropriate reaction to the mounting worldwide

squeeze on supplies of energy requires consideration of the CO₂ problem as an intrinsic part of any proposed policy on energy," the four scientists say.

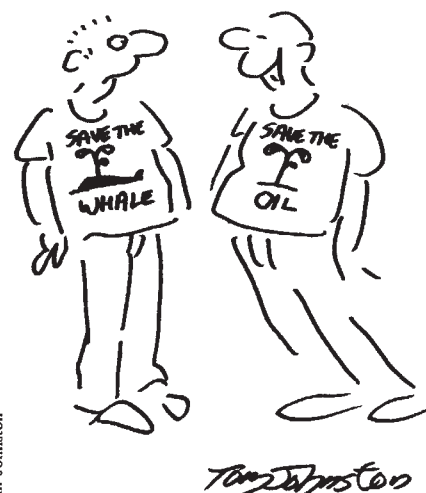
Environmental groups have warned the President about other potentially damaging aspects of a massive synthetics fuels programmes, ranging from the disruption caused by the construction of large production plants to the creation of large quantities of carcinogenic substances.

In a letter sent to the President last week, five such groups urged that he not allow himself to be caught up in what they described as a "synfuel panic", and consider lower-cost conservation programmes as an alternative to a massive commitment to synthetic fuels.

Many of the environmental risks associated with such a commitment were "unacceptable", the groups said in their letter; emphasis should rather be placed on more solar development programmes, less costly programmes for synfuel development, and more oil and gas exploration.

Support for such a policy also came from a different direction, namely a study prepared by the Harvard Business School, soon to be published by Random House under the title "Energy Future". According to this study, the need to constrain oil imports is much greater than most economists and energy analysts have realised, since the political and other costs of dependence on a few producing countries has to be included on top of the market price for oil.

The Harvard study suggests a diverse national energy policy encouraging modest growth in the use of coal, some reliance on nuclear power, and continued efforts to develop synthetic fuels; however it argues that in the current political climate, a stepped-up programme to accelerate conservation and solar energy would prove far more rewarding than any other energy path.



Tom Johnston

The environmental hazards associated with a major synthetic fuels programme were also covered by a report published last week by the Department of Energy's Office of Technology Impacts. Although the report says that it should be possible to find acceptable sites for large synthetic fuels production plants, it warns that successful commercialisation does have environmental risks, and that stringent environmental controls, management practices and "permitting conditions" should be maintained.

The Department report says that deployment of synthetic liquid facilities on an accelerated schedule to 1990 appears feasible in terms of current environmental constraints, although a target of 2 million barrels a day, rather than 500,000, would bring "rapid siting difficulties".

In reaching its conclusion, the report says there is some risk that environmental research and development programmes cannot fully satisfy all existing and expected regulatory demands, but that these risks should be known by 1985, and it is expected that appropriate control adjustments can be made.

"Reduction of these uncertainties requires refocusing of environmental research and assessment programmes to aggressively address these areas", the report says.

David Dickson

Congress slows growth of biomedical research

Appropriations committees in both the US House of Representatives and the Senate have recommended increases in the amount of money requested by President Carter for biomedical research in his budget proposals for the fiscal year 1980, submitted to Congress at the beginning of the year. However the increases, which average about 6% over the amount spent by the US on such research in the current financial year, is less than the expected rate of inflation — and considerably lower than the 24% increase over the President's proposal voted by Congress last year.

In his budget message, the President proposed that the research budget for the National Institutes of Health should remain virtually static, at about \$3.2 billion, between 1979 and 1980. Of particular concern to many research workers had been the proposal to reduce from 5,673 to 3,062 the number of new grants and competing renewals for investigator-initiated research (implying a drop from 42 to 20% of those approved for

funding actually receiving support).

On the advice of its appropriations committee, the House has recommended that the NIH budget be increased next year by \$202 million — or 6.4% over the current year; the major proportion of this increase will go to support investigator-initiated research in the various institutes, with the aim of increasing the number of new grants to 5,000.

"Thus the bill continues the real growth in federal support for basic biomedical research which has prevailed in recent years", the committee says in its report, adding that its primary objective "is to allocate funds to the most promising research endeavours as determined by the peer review system".

The committee also says that the present balance between non-competing continuations, competing renewals, and new grants "warrants a careful review to determine whether current procedures and mechanisms leave enough room for the support of unconventional ideas and new

investigators without proven 'track-records' ". It has requested the director of NIH, Dr Donald Fredrickson, to undertake such a review and to make a report before 1 October.

It also suggests that the NIH should explore the possibility of starting a programme to awarding stipends to medical or dental students to spend brief periods engaged in health research projects, with the hope of attracting them to research careers and thus increase the number of MDs and DMDs going into basic and clinical research.

The Senate is yet to vote on appropriations for biomedical research. However the Senate's appropriations committee last week accepted the proposal of its subcommittee that would cut \$13 million of the increase proposed by the House, still leaving intact an increase of 6% over the current year's funding. A single figure has to be agreed by both the House and the Senate before the appropriation can become law. □

Germany's research minister calls for 'critical science'

WEST GERMANY wants its scientists to be more critical in their research and to give due consideration to the disadvantages of technological advances as well as the advantages. There should be more discussion with pressure groups and the public before fundamental decisions on, say, nuclear energy are made.

This change of emphasis in Bonn's research and technology policy was outlined by Research Minister Volker Hauff when he released a major directory of research in West Germany recently. The 650-page report — *Bundesbericht Forschung VI* — is the Bonn government's sixth assessment of research policy, institutions and financing.

Hauff said the risks of research and technical progress should enter into decisions more than before. The government wanted to be more critical, to analyse and take account of the advantages and the disadvantages of new developments and to hold a dialogue with trade unions, individual population groups and pressure groups *before* taking decisions, not afterwards.

There is now a growing protest against almost every new large-scale technical construction in West Germany, even down to pressure groups against new roads and railway lines. Among the fields whose disadvantages need more critical examination, Hauff mentioned genetic engineering, the investigation of new sources of energy, and the widespread introduction of microelectronics. On the other hand, the research report, despite all the criticisms and objections surrounding Gorleben, still stands by the government's controversial "integrated disposal" policy

for getting rid of the waste from nuclear power stations.

The report records a total expenditure of DM30,000 million for research and development last year, of which half came from the state and half came from industry. It amounted to a 2.3% share of the gross national product. The main focus of R&D expenditure by the federal government (totalling DM8,500 million) was on the Federal Ministry of Research and Technology's promotion programmes (DM4,700 million) and on the Defence Ministry's research expenditure (DM1,700 million). The Länder on the other hand finance research mainly in the universities and other colleges.

The growth rate of the total R&D expenditure in the Federal Republic has declined from 20% per year in 1969 to 8% in 1978.

Within the R&D expenditure, the proportion for defence research and other government departmental research is declining, while that of the Ministry of Research and Technology's promotion programmes is rising slightly. The proportion of basic research fell from 19.5% in 1975 to 16.5% in 1979, but should rise again slightly to 17.1% by 1981.

The report also explicitly catalogues the aims of research and technology policy. In addition to the earlier three aims — expansion of scientific knowledge, increase in economic efficiency and improvement in citizens' living and working conditions — there are now a further two: the conservation of resources, and the "improvement of knowledge about technological risk", to provide a better basis for decision making. As far as

possible, all five aims are to be pursued with equal vigour. The inference is that the previously dominating emphasis on economic interests is to yield gradually in favour of a broader outlook. To what extent the non-economic aims stand a chance of realisation is, however, questionable.

This is evident in the Federal government's stress on how little latitude there is for 'going it alone' in the face of international competition, and in the increasing recognition that national decision-making mechanisms set narrow limits on its research and technology policy. While Research Report V was still saying briskly in 1975: "Research policy must investigate needs, evaluate various possible solutions . . . and set priorities"; this time the language is more reserved: "In an organisation of research which has, for good reasons, been designed pluralistically, government research promotion can and may lay emphasis on matters of content within a limited compass only."

Klaus Höpfner

●According to their annual report for 1978, just published, the Deutsche Forschungsgemeinschaft (which funds much of Germany's research in universities) spent over 700 million DM last year for more than 10,000 projects in all branches of science and research. The biosciences shared the largest fraction of the money, although their share sank from 37.1% (1977) to 36.2% (1978). The other natural sciences augmented their share slightly to 24.7%, while engineering sciences fall 0.5% to 20.9%. The most rapid increase came in the social sciences, which grew 17% from 1977 to 1978. □



'Spartan' scientist must go, says UK

Dr Tetz Yoshimura (above), the Japanese mathematical physicist, who seeks to stay in Britain and pursue his research undisturbed by the 'rat race' of modern Japanese life, is to be deported from Britain. His appeal against deportation was rejected last week by the Home Office on the grounds that his position "presents insufficient unusual features to justify giving him favourable treatment".

This represents a shift in standpoint from previous Home Office notifications, which rejected his application to stay on the grounds that his income from freelance coaching and translation work was insufficient to maintain him. Although Dr Yoshimura's spartan life-style allows him to maintain himself on £12 per week (half his average income), further guarantees were required. When the author Russell Hoban came forward as financial guarantor, the Home Office dropped the financial aspect, and turned to the fact that Dr Yoshimura's various academic commitments had terminated — not surprisingly, since coaching and the marking of examination papers is not normally required during the long vacation.

It seems strange that the Home Office claims the case exhibits no unusual features. Dr Yoshimura is an example of a now almost extinct race of scientists who hope to work without the support of major public funding bodies. When he came to Britain in 1971, he hoped to find a haven of academic detachment. Even the bulky file of correspondence on his case contains the occasional off-print or letter from an academic publisher about the proper placing of a bracket, that have somehow crept into the official hassling.

Faced with the final deportation order, Lord Avebury, chairman of the inter-party parliamentary Human Rights Group told *Nature* that he hoped that Dr Yoshimura's scientific colleagues would bring all possible pressure on the Home Secretary to reverse his decision.

Controls urged on mineral fibre

A unique government-industry-trade union working party organised by the UK Health and Safety Executive has proposed controls on mineral insulation wools used as substitutes for asbestos, and called for a programme of research to define further the hazards of fibres.

The report estimates that 15,000 workers are involved in the production and steady use of man-made mineral fibres (MMMFs) with up to 250,000 workers coming into "occasional" contact with MMMFs. In addition, the Chairman of the Party, D J Evans estimates that perhaps 20 times this number handle the materials privately.

The fibres are produced from glass, rock, slag and metallic oxides and include fibreglass and rock and slag wools which are widely used for home insulation as a substitute for asbestos. The proposed limits are a fibre limit of five fibres/ml and a mass limit of 5 mg/m³. (The present limits for asbestos exposure are 2 fibres/ml which is expected to be lowered to 1 fibre/ml shortly. The National Institute for Occupational Safety and Health in the US has proposed a new limit of 0.1 fibre/ml.) The time for introduction of the MMMF standard is "not to exceed three years".

In separate appendices, the trade union and industry representatives express the fundamental conflict in hazards legislation over who is to pay — industry in the form of safer equipment, or the work force in terms of poor health and possibly early death. Industry contends that no hazard has been shown to exist and as such controls must be based strictly on cost. "In the absence of a scientific basis upon which any changes in the (general dust standard) can be made, any change must be based upon what is considered to be the best practicable engineering controls."

Industry is willing however to support the introduction of a standard. "While emphasising the absence of a scientific basis for the figure we accept the proposed standard of 5 fibres/ml." This standard is not expected to affect industry operations significantly.

In their statement the TUC states that it would prefer a standard of 1 fibre/ml and feels it could live with a standard of 3 fibres/ml as "an acceptable and attainable standard under present engineering control knowledge." The TUC disagrees with the 5 fibre/ml standard and it also disagrees with the three year lead-in period, calling instead for adoption of the standards within 12 months.

In the middle of the political fight between those who assign the limits, and those who bear the risks, stand the scientific specialists. In the UK, research into the hazards of fibreglass is being conducted at three places. In Southampton, Professor Donald Acheson will conduct human epidemiological studies. In Edinburgh, the Institute of Occupational Medicine is working on standardisation techniques, and at Penarth a Medical Research Council unit will conduct animal studies as well as designing measurement devices. Additional research is being conducted by the industry funded. International Agency for Research into Cancer with support of the World Health Organisation.

Experienced health and safety observers, however, point out that the situation bears an uncomfortable resemblance to the history of asbestos. In the UK, the first cases of asbestosis were reported by lay people in 1898, and the government noted cases of death in 1900. But the first legislation was not introduced until 1931, and widespread public awareness of the extent of the hazard did not occur until the 1970s. Fibreglass and its cousins are relatively recent materials and the historical data may not even exist.

Dr Morris Greenberg, medical adviser to the Health and Safety Executive thinks that the situation is difficult. "We think that MMMFs are not in the same league as asbestos but we don't know how far down the league table they are." **Joe Schwartz** *Man-made Mineral Fibres*. Report of a Working Party to the Advisory Committee on Toxic Substances. HMSO. 50p.

Asbestos, not glass fibre, killed woman

Scientists at the US National Institutes of Health announced last week that a case of pneumoconiosis reported by Japanese research workers last autumn as possibly being related to exposure to glass fibre had turned out to be the result of previous exposure to asbestos. The link between glass fibre and pneumoconiosis had been suggested by a Japanese research worker at a meeting last year held under the auspices of the National Cancer Institute, following his discovery of physiological damage in the lungs of a woman who had worked with glass wool similar to the damage caused by asbestos fibres.

The report received considerable publicity, and generated wide concern in the glass fibre industry. However the NIH reported last week that microscopical analysis of biopsy material carried out by Dr I J Sellikoff of the Mount Sinai School of Medicine showed the presence in the lung of asbestos fibres, but not glass fibres.

The NIH report says that during discussion of the original report, it was understood that six other cases of pulmonary disease had been found. "The existence of these cases cannot be confirmed and is attributed to a linguistic error", the NIH says.

news in brief

IWC offers some hope for sperm whales: Conservationists won a half-victory at last week's International Whaling Commission meeting after years of unsuccessful campaigning. There will now be an indefinite ban on factory-ship whaling for sperm whales, and a whale sanctuary will be set up in the Indian Ocean for ten years. About 7,000 sperm whales will be saved by the factory-ship whaling ban, which will probably put an end to the Russian whaling fleet. The small minke whale, however, has not received the same protection. Commercial minke whaling has been excluded from the IWC moratorium, and the Antarctic end of the Indian Ocean where the minke is found in great numbers will not form part of the sanctuary.

The IWC also set a new quota of 16,000, some 4,000 less than last season's kill. Recommendations were made by IWC's technical committee on humane methods for slaughtering whales but no nation offered to allow independent observers on to their ships, so research on death times and killing methods will be difficult. The IWC also agreed that scientific permits and quotas on the number of whales needed for research would now have to be approved by a science committee. Previously each nation set their own quotas.

Conservationists were disappointed that a proposal to ban all sperm whaling was rejected by the IWC and that shore-based sperm whaling will continue. The commission also rejected a proposal to prohibit trade between IWC and non-IWC nations, and a proposal to add the narwhal and white whales to the IWC's list of protected species. The commission also failed to come to a decision on pirate whaling.

Malformations at birth increased in Séveso last year: Figures issued by the special regional commissioner for Séveso, Mr Antonio Spallino, suggest that there has been an increase in the past year in the number of local children born malformed. According to Spallino, of 2,777 births recorded in the Séveso region in 1978, 90 infants were born malformed and 24 were classed as 'doubtful cases'. The comparable figures for 1977 were 2,756 births, 41 malformations and 25 doubtful cases. In July 1976, a trichlorophenol reactor belonging to the Givaudan company at Séveso, overheated, discharging its contents — including a teratogen 2,3,7,8 — tetrachloro-dibenzo-p-dioxin (dioxin) — over a populated area near Séveso in Italy. Concern was expressed at the time that the presence of dioxin, a well-known teratogen, could lead to an increase in birth malformations. The current figures do suggest such an increase, although the rate of malformations reported for 1978 is still comparable with the norm for Western Europe and bears little relation to the concentration of dioxin.

MRC staff committee criticises MRC management for budget cut inaction: Officers of the National Staff Side of the UK Medical Research Council have mounted a campaign to oppose cuts in the medical research budget, that includes strong criticism of MRC officials. The committee cites the closure of two MRC units and the future loss of 380 jobs or 10% of the workforce as evidence of the damage being done to the nation's research effort. "We do not agree that the Council should implement or even accept the imposed cuts. The MRC seem to have done nothing to offer resistance on their own account and have refused to be associated with our protests." Officials in the MRC claim that further assessment is needed and that management has its own separate channels of communication with government. "No figures yet exist for losses due to budget cuts and the two unit closures were for scientific not economic reasons" an MRC official told *Nature*. "The official side has made its views known to the government through the Advisory Board for the Research Councils." The Council's views have not been disclosed pending a Council meeting called for 19 July.

Aldermaston unions to file contamination lawsuits: Four trade unions at Aldermaston, the UK's nuclear weapon laboratory, are considering more than 70 claims for damages against the Ministry of Defence for plutonium contamination. The action follows the unions' refusal last February to work in contaminated areas identified by the Pochin report. The report was produced by Sir Edward Pochin, a radiologist, after plutonium had been found in the lungs of three workers in the laboratory's laundry room. Nine workers in the industry-chemistry section were also contaminated.

Sussex students now expelled: Sussex University students Richard Flint and Shaun Fensom have been expelled from the university for their part in preliminary science examination disruptions. The expulsion was the result of a second disciplinary committee action announced on 6 July. A faculty committee consisting of pro-vice Chancellor Professor T Nuttal as chairman and four faculty members, M Hawkins, W Lamont, L Cook, and J Maynard-Smith, appointed by the university Senate had reinstated the students on 25 June — provided they were "not found guilty of a further breach of discipline as defined by the disciplinary code after this judgement". One week later the committee reconvened to consider the student disruptions of the week of 18 June as a separate series of events from the first disruptions in May. The June demonstrations saw the breaking down of doors and the first arrests and injuries, when faculty members confronted a picket line of 150 students in front of the Building of Engineering and Applied Science. A faculty petition signed by over one hundred faculty members — few of them from the science faculty — called for the reinstatement of the two students (as well as a withdrawal of threatening letters to university student rent strikers (21 June)) to no avail.

In addition criminal charges of malicious intent to cause damage have been filed against five students and the university has taken out a £20,000 order of compensation against five students and a £10,000 order of compensation against the student union.

UK government scientists win claim; strike goes on. Scientists of the striking Institution of Professional Civil Servants have won their pay claim for this year. The Civil Service Department has granted pay rises in five scientific grades ranging down from 34.1% for the top grade of Principal Scientific Officer to 22% for Assistant Scientific Officer. The rises are linked to those for administrative grades — a hotly contested point. In what has become an increasingly bitter series of negotiations the IPCS has also won its demand that next year's salaries are to be negotiable. The CSD had demanded the union accept without negotiation the results of a pay research exercise to determine their pay.

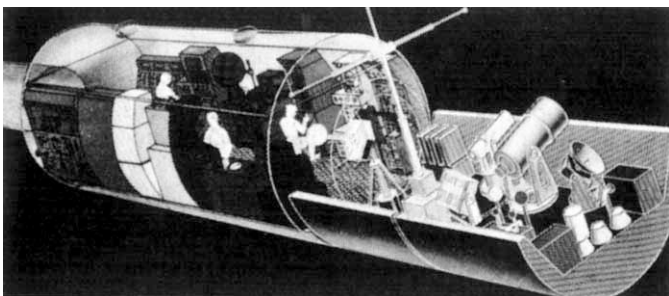
In the meantime the IPCS is still on strike in support of claims for the governments 40,000 technologists — an action which is supported strongly by the scientists. Government scientists at Royal Ordnance factories at Glascoed, South Wales and Bradeway Green and Chorley both near Manchester have brought shell manufacture to a halt. In addition scientists at the government's Forensic Laboratory withdrew their labour in a two week strike but, because of the Official Secrets Act, could not disclose the effect of their action even to the union leadership. The IPCS has called for arbitration and is continuing its tactics of selective strikes, half day strike-demonstrations and withdrawal of good will.

NASA orders a Spacelab from ESA: On 5 July, NASA signed a contract with the European Space Agency for the supply of components for a second Spacelab flight unit identical to those of the first Spacelab unit which ESA is currently building. The cost to NASA of the second Spacelab will be announced by ERNO, the

German head of the consortium which will be building it, at the end of this month. NASA hopes however, that it will not exceed 110 million accounting units (1 a.u. is \$1.25).

A memorandum of understanding that NASA should buy a second Spacelab from ESA was signed when the two agencies first embarked on the Spacelab programme. During much of last year, ESA and NASA spent time trying to think of ways in which NASA could procure the second Spacelab without any exchange of funds. A system of bartering, however, was unattractive to ESA which was reluctant to provide a Spacelab in exchange for free shuttle flights for some of the first Spacelab's demonstration missions. Difficulty in filling some of the demonstration missions with sufficient experiments, mainly because of costs, meant that ESA could not guarantee in advance that it would use the shuttle flights it had bartered for. The two agencies therefore finally agreed to revert to the normal way of selling and buying by exchange of funds.

Funding of the first Spacelab is currently undergoing some difficulties, however. Present estimates of the total cost are now 680 million accounting units, about 140% of the original estimated cost. Under the Spacelab agreement, any participating state is entitled to leave the programme when costs rise to 120% of the original estimate. In principle, however all states want to continue despite the 40% overrun — but some, notably Italy, are finding difficulty in raising the funds.



Another factor which could add further to costs is a possible further delay to the Spacelab launch caused by difficulties with the space shuttle. Despite the recent statement by Dr Robert Frosch, NASA's administrator, that the space shuttle's first operational flight will now be delayed until September 1981, however, ESA has received no formal notification of any further delay to the first Spacelab launch most recently scheduled for August 1981.

Religious leaders criticise links between knowledge and power:

Scientists and technologists have weakened their claims to intellectual freedom by showing bias on the side of those who wield economic power, according to Dr Philip Potter, general secretary of the World Council of Churches. Addressing the opening session of a two-week conference on science and religion in Boston last Thursday, Dr Potter said that there was little sign that scientists were on the side of the oppressed, the deprived and the "marginalised".

"Science and technology might claim to be objective and universal, but this claim is not born out in reality," said Dr Potter, who is a methodist minister in Jamaica. "Science and technology are not neutral or value-free but are instruments of power, and that means political power"; a leading issue of the conference was therefore how ways could be found of allowing the public a greater share of decisions on how science should be employed.

In contrast to Dr Potter's comments, Dr Robert H Brown, an astronomer from the University of Sydney, Australia, defended science against claims that it had been corrupted through colluding with the powerful against the interests of the poor. "Some tell us that the ideals of science have been so eroded by its alliance with power that it has little to offer us," he said. "I suggest that it is no more a valid reason for turning away from the values of science than it was for scientists to turn away from the Christian faith."

Laetrile claimed to increase size of tumours: Doses of the controversial drug Laetrile, which some claim to have anti-cancer properties, has been shown to increase the size of tumours in rats, and eventually lead to cyanide poisoning, according to the report of two US medical research workers in the *Journal of the American Medical Association*. The researchers, from the Northwestern University medical school, say that doses described as "realistic in terms of human ingestion" led to a progressive increase in the size of tumours when administered to rats, and as the dosage was increased, so to did the incidence of death from cyanide poisoning. "The findings seriously question the use of amygdalin [Laetrile] in clinical medicine under any circumstances" said Dr Janardan D Khandekar, one of the scientists involved.

Meanwhile the Food and Drug Administration has rejected the charge from one of its staff scientists that it is unethical for the agency to study the effects of Laetrile, as has been requested by the National Cancer Institute, even though there is no scientific evidence of its efficacy.

In reply to a citizen's petition filed by Dr Robert Young, the agency has said that it does not consider such studies to be unethical since the tests might provide some useful information. "The ethical aspects of failing to discover a useful cancer drug are complex and not well explored, but the prospect of such an outcome occurring because of intellectual rigidity or over adherence to an ethical dogma is unsettling," replied Dr Richard Crout, director of the Bureau of Drugs, although admitting to "a great deal of ambivalence" on the matter.

Germany's science council recommends sites for new polar research institute:

The federal government of West Germany, which aims to join the consultative round of the Antarctic Agreement, wants to strengthen polar research activities in the FRG and therefore decided some time ago to establish a central polar research institute "as soon as possible". Last month, the Science Council (Wissenschaftsrat), which was asked to advise, welcomed the initiative and made some detailed suggestions for the new 30-scientist establishment.

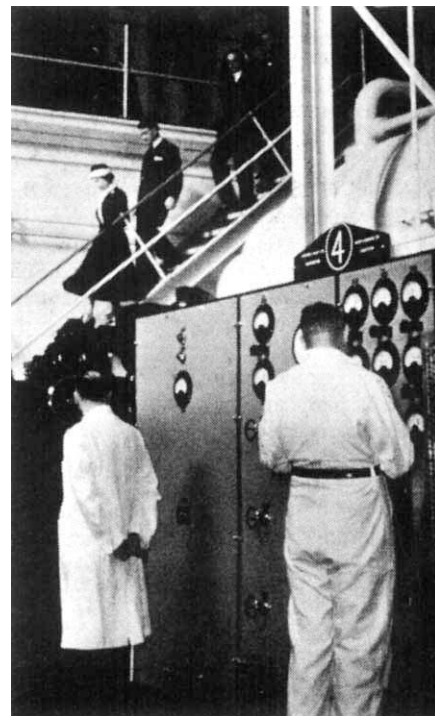
The Science Council recommends that the institute should set up as a foundation under public law and financed equally by the federal government (the *Bund*) and the *Land* where it is situated. It should be responsible for the planning, construction and operation of the German research station in the Antarctic and direct the operation of the ice-breaking research and supply ship in the polar regions. Moreover, it should have an important function in co-ordinating national and international polar research activities, thereby co-operating with other institutions acting in this field and building on a long-standing tradition of German polar research. Concerning its location, the Science Council recommends that it should be established in the neighbourhood of a large university with internationally proven research in the relevant fields. The Science Council therefore put Kiel at the top of the list, Hamburg next, and Bremen (including Bremerhafen) third. Among the three towns, Hamburg has rejected the institute, as it is fully occupied with establishing a new technical university. Kiel, on the other hand, is showing a strong interest, but faces strong political objections in Bonn.

India to launch oceanographic research vessel: An oceanographic research vessel will be commissioned by 1981 for research and development of marine resources in India. The Indian government's decision to acquire the vessel is in line with the government's recognition of the importance of ocean waters and the sea bed as store of living and nonliving resources. Earlier, the Government of India had enacted the Maritime Zone 1976 which reserved the right of survey and exploration of the marine resources of a 320 km 'Exclusive Economic Zone' and the continental shelf exclusively for the Indian people. To ensure optimum utilisation of research facilities and proper co-ordination the Ocean Science and Technology Agency (OSTA) was set up in 1978. The research vessel is to be manufactured in the Federal Republic of Germany.

UKAEA: boom or bust?

Two years after the establishment of the United Kingdom Atomic Energy Authority 25 years ago, the Queen was opening the world's first full-scale nuclear power station at Calder Hall (see right). But these early boom years of British nuclear power are over. Energy forecasts are drastically lower, and uncertainty about the future of the industry has led to the loss of key staff. The fast breeder design team, for example, has been virtually disbanded. So even if the new government wished to re-invest in the nuclear industry — as the Prime Minister in-

dicated at the recent summit meeting in Japan — it would first have to re-establish confidence in the industry itself. So while Britain's new energy ministers, headed by Secretary of State David Howell, consider the formation of a new energy policy for Britain, we asked **John Surrey** of the Science Policy Research Unit at the University of Sussex to give us his view of the future of nuclear power in this country. On a following page, **Paul McDonald** considers the difficulties facing another major energy option — conservation.



How the UKAEA might survive

LONG-term forecasts given at the UK National Energy Conference in June 1976 pointed to a future 'energy gap', due to rapid growth in demand and rapid depletion of British North Sea oil and gas reserves. They indicated that without major investment in coal and especially nuclear power, the UK would have to rely on increasingly large and expensive oil imports after about 1990. Only the gas industry challenged the forecasts, arguing that North Sea supplies would last until at least 2000. For the other fuel industries, high demand forecasts promised security and expansion.

Coal's target capacity for 2000 was set at 170 million tonnes — the maximum attainable given the need to replace a large proportion of current capacity and the long lead times for major projects such as the development of the newly discovered coal-fields at Selby and Belvoir. The same climate encouraged the UKAEA (Atomic Energy Authority) to present a 'reference' programme which envisaged 104,000 MW of nuclear power in operation by 2000, including 33,000 MW of fast reactors. These were to be preceded by a 1,300 MW demonstration plant, originally code-named CFR-1 and later CDFR. The 'reference' programme would require about 6,000 MW of nuclear plant to be built *each year* — somewhat more than the total nuclear generating capacity in service today.

A recent consultative document, the 1978 Green Paper on energy policy, contained much lower demand estimates for 2000 — 450–560 mtce (million tonnes of coal equivalent). For comparison, the mid-point of the 1976 forecasts was 558 mtce and the top of the range was 760 mtce. The 1978 official forecasts were very close in several respects to those pub-

lished by Chesshire and me shortly before the Green Paper appeared. The two sets of estimates were very close for final energy consumption and, in the 'high' case, for primary energy. But in the 'low' case the Green Paper estimate of primary energy demand was 115 mtce higher than ours (see Table 1). Given the other similarities, this big difference can only stem from an assumption behind the official forecasts that electricity consumption, and therefore conversion losses, will be much greater than we estimated under the 'low' case. The discrepancy simply illustrates that the prospects for electricity are one of the biggest uncertainties in energy planning — which is of key importance for coal and nuclear power, given their dependence upon electricity generation.

Until nuclear power stations can work satisfactorily and economically under highly variable load conditions, the scope for nuclear power will be limited to base load generation. In the UK, it is further restricted by a large plant surplus, including a large amount of plant still under construction. Furthermore, electricity is expensive relative to other fuels: in the industrial market the price ratio of electricity is 5.4 relative to gas, 6.1 relative to coal and 3.8 relative to oil.

As long as the relative price is so high, electricity will not substitute for oil and coal in competitive crude heat uses, or for gas in many premium uses. Even assuming that nuclear-generated electricity will henceforth be cheap relative to oil and coal, the fact that nuclear power currently

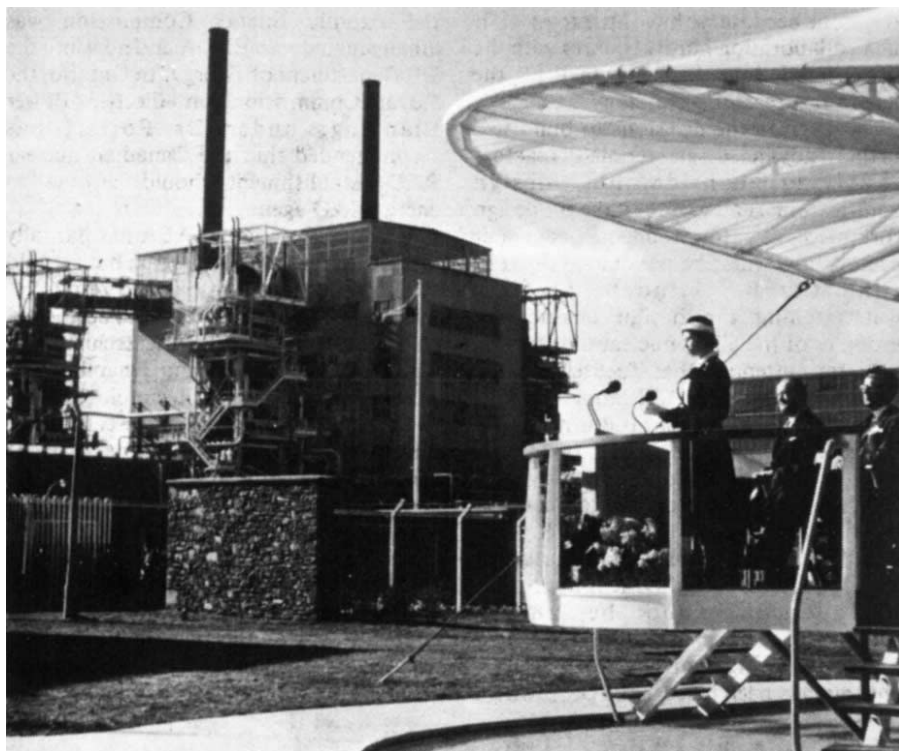
Table 1 1978 Forecasts of UK Energy Consumption in 2000

	1977 Actual	Forecasts for 2000 SPRU ¹ billion therms	Green Paper ²
Domestic	15.0	14.2–16.2	(13.6)–15.8
Transport	13.1	17.2–19.9	(15.8)–17.0
Other final consumers	7.6	6.0–8.7	(7.0)–9.0
Iron and steel	4.9	3.7–8.4	(7.0)–9.9
Manufacturing industry	17.9	19.6–28.2	(20.0)–25.4
FINAL ENERGY	58.5	60.8–81.4	(63.4)–76.8
million tonnes of coal equivalent			
Coal	122.7	83–162	170 ³
Oil	136.6	124–224	150
Natural gas	62.8	88	50–90
Nuclear and hydro	14.3	40–102	95
PRIMARY ENERGY	338.4	335–577	(450)–560

¹ J. H. Chesshire and A. J. Surrey, *Estimating UK Energy Demand for the Year 2000: A Sectoral Approach*, SPRU Occasional Paper No. 5, February 1978.

² *Energy Policy: A Consultative Document*, Cmd 7101, HMSO, February 1978. The figures in brackets apply to a 'low' case which is only partly specified.

³ These are estimated availabilities of indigenous fuels only. They sum to 465 mtce, in the 'high' case, leaving net imports of 45 mtce and a 'policy gap' of 50 mtce.



accounts for only 13% of electricity supplied to the British grid means that the price of electricity will not fall appreciably relative to the price of oil and coal until base load is supplied predominantly by nuclear power stations. That, of course, would involve a large nuclear programme to replace the fossil fuel power stations built in the 1950s and 1960s and — since two-thirds of British coal output goes to power stations — a big fall in coal demand.

One is inescapably drawn to the conclusion that electricity demand growth over the next two decades will continue to be low, especially if the supply of North Sea gas reaches 6,000 million cubic feet a day in the early 1980s and remains at that level until 2000. In these circumstances, electricity demand growth will be chiefly in certain premium uses where electricity offers advantages which outweigh its high price. Information available on domestic appliance ownership levels and industrial process applications indicates that the growth potential from these electricity-specific uses is quite small.

Compared with the earlier 104,000 MW 'reference' programme the 1978 Green Paper sees a nuclear component of 25-40,000 MW in 2000, implying the construction of only 1-2,000 MW annually from 1985. The central question now facing the nuclear industry is no longer whether it can muster the skills and resources to build a large programme, but whether it can survive a long period of low home ordering and continued difficulty in exporting.

Following Ince B, Drax B and two new advanced gas cooled reactor orders, further advance orders for power stations should be out of the question on cost grounds. It is therefore urgent to identify

the resources and skills that are specific to nuclear design and engineering and to examine how they can best be organised and retained against the time that series ordering can take place. This problem must have precedence over the fast reactor question, for it would be sheer folly to develop fast reactor technology without ensuring that the design and manufacturing base is there so that nuclear power stations can be built efficiently when they are needed. If there is a solution, it is likely to require the rationalisation of the whole power plant manufacturing industry, including nuclear design and construction and large steam turbine generators.

The recognition that long lead times make it impossible to install a significant fast reactor programme by 2000 now seems to have brought a corresponding change in the reason adduced for proceeding with the demonstration plant. When CFR-1 was regarded as the first of a large fast reactor programme, it could be held that costs of the demonstration plant over and above the costs of the cheapest alternative power source (a light water reactor, for example) would be recouped from the benefits accruing from the commercial fast reactors that were soon to follow.

Since the justification was primarily economic, it was possible to argue that CFR-1 would carry no opportunity cost in terms of other R&D projects. Questions about resource allocation were answered by the claim that the programme as a whole would show a large and positive net present value after applying the test rate of discount that is applied to public sector investments. As long as it appeared tenable to argue that the cost of CFR-1 would be recouped from the subsequent commercial

programme, the relevant economic questions concerned the estimated capital and operating costs of fast reactors compared with thermal reactors and the estimated long-run supply and price of uranium. Both questions involve major uncertainties and evoke widely differing views.

If, however, approval for CDFR is seen as carrying no commitment to build a programme of fast reactors before the turn of the century, the choice must be seen primarily in the context of R&D strategy. If the commitment which is sought is for CDFR alone, not for a subsequent programme, the question of R&D opportunity costs looms large — the resources must come from somewhere and something will have to be foregone. Unless it is reasonably certain that a highly economic programme will follow in the not-too-distant future, the question of economic risk becomes very important with an R&D project costing £1,500 million or more. Such a commitment to any single R&D project can scarcely fail to affect the climate of opinion and the funds available for alternative R&D options. We therefore need to be quite sure that an equivalent funding of alternative technologies (not only in the energy field, of course) will be less rewarding than building CDFR.

The argument that Britain should acquire the capability to build commercial fast reactors as an insurance against long-term uranium scarcity involves a political judgement. It requires a decision on the acceptability of the whole range of risks and the appropriate rate of social time preference, how much the present generation is prepared to forego for the well-being of future generations.

Another argument in favour of CDFR, the 'export' justification, is wholly spurious for no one can say whether there will be opportunities 20 years hence to export 1,300 MW fast reactors. So far, domestic markets in the USA, Japan and most of Western Europe have been closed to imports of nuclear reactors. Whether the developing countries can accommodate 1,300 MW generating units on their power systems is a moot point. And on foreign policy grounds it is surely unwise for the UK to appear willing to promote exports of fast reactors (and the plutonium to fuel them) before rigorous international safeguards for reprocessing and the commercial use of plutonium are agreed and implemented.

If the political judgement is that Britain should acquire the capability to build fast reactors in the future, the relevant question then is how to proceed. Several further questions must then be answered, because building CDFR is only one of the options for acquiring this capability.

Firstly, why is it necessary to scale up to 1,300 MW? Doesn't the ability to replicate and perhaps stretch the 250 MW Dounreay prototype fast reactor give sufficient insurance against the risk of long-term

uranium scarcity? If there are compelling technical reasons for scaling up, why not concentrate R&D effort upon the specific, critical metallurgical and engineering problems, as opposed to building a 1,300 MW power station costing at least £1,500 million? In any case, the performance of very large conventional and nuclear generating units hardly inspires confidence that the proposed scaling-up is economically justified.

Second, since France and Germany are intent upon building demonstration plants which are broadly similar to CDFR, why not wait and license from them when their designs are proven? After all, other countries have acquired the ability to build thermal nuclear reactors on the basis of foreign licences, whilst Britain has spent untold sums in developing indigenous reactor designs. Alternatively, why not collaborate with France and Germany, learning through participation, so as to be able to incorporate any necessary modifications in an eventual British version? The argument that Britain must

go it alone because we have little to offer in such collaboration hardly squares with the oft-repeated claim that Britain leads the field in fast reactor technology.

Third, given the decisions to build two further advanced gas cooled reactors (involving considerable design modifications) and to undertake the design work necessary for a pressurised water reactor, how will it be possible to proceed with CDFR without severely overstressing the design engineering resources of the ailing nuclear industry?

So far, attention has focused on the question of whether CDFR should be built, rather than what happens if it is not built. Fast reactor work accounts for two-thirds of the scientists and engineers employed by the UKAEA on reactor development. A decision not to build CDFR will immediately call into question the future of the UKAEA. Even if CDFR is approved, this question cannot be ignored indefinitely. Sooner or later, it must confront any single-mission R&D agency. It has already been faced in the USA, where

the Atomic Energy Commission was amalgamated into ERDA and now into the US Department of Energy. In Ontario, the Royal Commission on Electric Power Planning, under Dr Porter, has recommended that the Canadian nuclear R&D establishment should become an energy R&D agency.

It is true that the UKAEA has partially diversified into ancillary work; but I would like to think that its skills and resources, which are unique in Britain, can be applied on a much larger scale to other technologies with the same brilliance and dynamism that were applied to nuclear technology in the 1950s. This is not to suggest that the UKAEA must become an energy R&D agency. Its new role must be defined in the context of overall industrial and R&D strategy and the need to make the best use of all our public laboratories.

The assurance of a future rôle which is nationally and personally worthwhile will remove much of the uncertainty for the scientists and will make it less likely that a breeder is built for the wrong reason. □

'Saving it' is easier said than done

With the oil crisis in full cry and nuclear energy still an uncertain option, the UK government this week called on industrialists to make energy saving a priority. **Paul McDonald** assesses the potential of this new policy.

Energy conservation, according to Mr David Howell, the Secretary of State for Energy, is "now at the centre of energy policy". Faced with the task of trying to cut oil consumption by five per cent, the Government is relying heavily on industry to make significant improvements in the efficiency with which it uses its energy. However, apart from the widely publicised efforts of certain individual companies, there has been little progress in this sphere.

There are no major technical reasons why much of industry could not improve the efficiency of its use of energy substantially. Such is the inefficiency of the average industry, that its energy can usually be reduced by 10% with little or even no capital investment, simply by goodhousekeeping measures, such as insulation, or turning off machinery not in use. Savings of a further 10% are usually possible from improvements in processes; and a further 10% on top of that from major reorganizations of processing and wholesale re-equipping. The actual proportion that can be saved varies from industry to industry. The Department of Energy's estimates of potential savings vary from a total of 10% in the pottery industry to 61% in aluminium smelting.

The Department of Energy has, unfortunately, no way of assessing accurately how industry is progressing in

this. The most optimistic estimates put the average level of improvement in the efficiency of industrial energy consumption since 1974 at 10%; but this is no more than could be achieved by general goodhousekeeping measures in most industries — and certainly below the level of our overseas competitors such as West Germany and Japan. In spite of an upsurge of interest in energy conservation amongst industrialists since January, there are still many factors inhibiting any major improvement in the use of energy, which are going to make the Department's target of 5% very difficult to achieve.

On the surface, there appears to be a good deal of activity: over 4,000 energy managers have been appointed; but their functions and responsibilities vary enormously, and, even in some large companies, a number of them are little more than token appointments.

A major constraint on improving energy efficiency is the lack of capital for measures designed to improve energy use, such as waste heat recovery and improving processes. This is particularly the case for the large number of industries where energy accounts for less than about 2% of manufacturing costs. There is, in any case, a preference in most industries to invest in production rather than conservation. Whilst the latter brings savings, these do not necessarily show up in reduced expenditure on energy in subsequent years, since they may easily be cancelled out by the general rise in energy prices. Given the preference for increasing output, conservation measures are often penalised by the imposition of shorter payback times than those for other items of capital expenditure.

Continuing economic recession and low levels of productivity greatly inhibit the efficient use of energy by forcing industries to work well below their full capacity. Furthermore, in a number of industries, such as food, drink and tobacco, there is a trend towards increased processing, which will increase the consumption of energy for the same level of output.

Improvements in energy use are further hindered by the slow rate of re-equipping in British industry. Many companies wait until equipment is due for rebuilding or replacement before trying to improve its energy consumption. In the case of a furnace, this is only once every 5-10 years.

Proper records of fuel consumption are vital to the success of any serious attempt to improve energy efficiency. In a great many companies these simply do not exist, making it impossible to set targets or monitor progress. Furthermore, the high costs of installing and maintaining adequate metering for this task will make the accurate control of energy use a continuing problem for a large number of firms.

The aim of company energy policies is not always to save energy. For certain firms operating continuous processes, security of supply may be the paramount consideration. For others, energy policy may be part of a general cost cutting exercise, and can just as easily involve renegotiating tariffs or switching to other fuels, as actually trying to reduce consumption by requiring the installation of energy using anti-pollution equipment as part of its health and safety policies. Energy conservation is clearly not going to be the easy part of the Government's energy policy.

Dr McDonald is a writer on energy matters

A political view of CO₂

The increase in atmospheric carbon dioxide may be accelerated by President Carter's new-found enthusiasm for synthetic fuel. But the atmospheric 'crisis' may come too slowly to bother the politicians, argues **Michael Glantz**.

FOR those interested in the study of climate — including its impact on society and society's impact on it — 1972 was an extremely important year. In that year a collection of weather anomalies occurred adversely affecting global food production and therefore availability. At that time some blamed the food shortages on the weather. More recently, however, those claims have been reevaluated and the blame is being apportioned more correctly between weather and society. The anomalies of 1972 included the fourth consecutive year of drought in the Sahelian zone of West Africa, the failure of the Peruvian coastal fisheries, droughts in Central America, the Soviet Union, India and China, along with excessive rains in parts of the Philippines, Australia and Kenya.

At that time the sharp increase in grain prices on the international market, the largest grain purchase by the USSR to that date, and the general scarcity of foodstuffs in the international marketplace, were all blamed on fluctuations in the weather and led to the concern that the weather had gone haywire. Concern about what was happening to the weather can be evidenced by citing such publications as "Ominous changes in the weather," "What's gone wrong with the weather?", "When the Sahel freezes over," The Cooling, Blizzard, Ice, The Genesis Strategy, Hot House Earth, Climatic Change, The Weather Conspiracy and so on.

Was the climate regime getting cooler? Was it getting warmer? Was it staying the same? Was climate variability increasing? Climate (and weather) have become recognised as important variables in the activities of societies, regardless of the level of industrial development, geographic location, or ideological bias of the respective society. This new-found importance is due partly to the media's acceptance that such factors are newsworthy, and partly as a result of the raised consciousness of some political and economic leaders that climate factors must be considered in the food-population equation.

A projected CO₂-induced global warming became part of the debate and is increasingly being viewed as an important



scientific and socio-economic problem. In addition to the natural processes, CO₂ enters the atmosphere as a result of human activities, primarily the burning of fossil fuels. The current rate of increase of CO₂ content in the atmosphere has been measured at about several tenths of one per cent per year. With projected energy use increases it has been suggested that a doubling of the CO₂ content in the atmosphere can be expected by the middle of the 21st century.

A doubling is expected to warm the lower atmosphere (as a result of various physical processes) with estimates ranging from 1 to 4 °C.* This warming is expected to affect weather patterns and climate regimes around the globe, changing precipitation and temperature patterns resulting in the shifting of agricultural zones. Such shifts would adversely affect some regions while favourably affecting others, with no assurances as to who the winners and losers will be. Thus, much attention has been focused on the ultimate effects of the increased CO₂ loading of the atmosphere.

In fact, the increase in the CO₂ content of the atmosphere takes place each day

and, measurably so, each year. Yet, the annual increments are perceived by most observers to present little immediate danger to society and are considered non-threatening. No immediate attention, let alone rectification, of the CO₂ problem is perceived to be necessary.

The CO₂ impacts are especially an important area of study for the social sciences because they represent, as well as highlight, a relatively new (some would say neglected) research area of interactions between climate and society. Many consider the study of the CO₂ content of the atmosphere timely even though it is still only a potential problem with what appears to be a relatively long lead time, that is, the major impacts are projected in terms of decades. The "luxury" of a long lead time is not accepted as such by everyone, since society will require a long time to make adjustments that would be in line with attempts to develop alternative sources of energy to reduce the use of fossil fuels.

It may be useful to separate the CO₂ problem in to the two predominant perspectives: as an event — the doubling, and as a process — the annual 1 p.p.m. increase. In reality both views are correct but each by itself presents an incomplete picture of the problem. To better understand and, therefore, deal with the potential societal and environmental

**It is important to keep in mind that the locations, types, magnitudes and directions of change as well as the importance of certain variables in CO₂-related processes are in most instances educated speculation at this time.*

problems associated with an increased CO₂ content of the atmosphere, the problem must be viewed as both event and process.

Most of the attention of the media, the public and the policy-makers is directed toward the CO₂ problem as an event because of the spectacular aspects of a CO₂ doubling. For example, an often referred to impact of a doubling is a surging into the ocean of the West Antarctic Ice Sheets resulting eventually in an increase in sea level on the order of 5-8 metres.

Spectacular impact of other natural disasters

The CO₂ problem viewed as an event has been likened to other specific natural disasters — earthquakes, droughts, floods, severe or dry winters. The impact of a drought, for example, on the environment and society can often be spectacular. In many instances, however, the drought only highlights a process of environmental degradation that has been under way for a period of time preceding the drought event itself; a process which at any one point in time represents only a low level abuse of the environment.

So, the dust storms in the American Great Plains during the droughts in the 1930s and briefly in the 1970s, while spectacular, were not unique (except for their magnitude). Each year during the period between such major drought events, the top soil was also blown away at a much less spectacular level. The cumulative effect of such storms in the interdrought periods may in fact be much greater than that of the specific drought event.

One might argue by analogy that the tendency to overfocus on the more spectacular event (the CO₂ doubling) takes place at the expense of attention to the slow but steady annual increases in CO₂ (the process).

The example of urban air pollution, as a low-level but continually increasing insult to the environment, in some ways represents a microcosm of the CO₂ problem as a process. It also suggests how society might react to such a problem. There are very gradual, but definite, changes that take place in the environment (especially weather and climate) as a result of urban air pollution, a large part of which results from the burning of fossil fuels.

Yet, in many instances the changes that have occurred are low level and have apparently taken place gradually over an extended period of time. One might ask, how well have societies and their political leaders dealt with those incremental changes in air pollution levels and with the social changes they may have brought about? How well have they dealt with drought episodes? How well can they be expected to deal with CO₂ problems?

Studies of societal responses to

spectacular events such as major droughts and their impacts suggest that political interest in those kinds of problems varies with the presence of other problems that compete for time and scarce resources. In addition, interest in such events also rises and falls as a result of the different lengths in office of different decision-makers (2,4,6 year cycles for politicians). This fluctuation in interest is manifested in a relatively abrupt swing between two types of decision-making processes:

- "crisis management" and
- "muddling through."

In crisis decision-making, the time to act is perceived to be short, the stakes are perceived to be high, and the impacts are seen to be severe. As for "muddling through," the opposite is seen to be the case; there is perceived to be plenty of time to act, society is expected to rise to the occasion with a swift and appropriate response if and when it becomes necessary, and the impact of the next situation is expected to be unlike the last one because of the uniqueness of meteorological events (or because of the belief that society tends to learn from its past mistakes), or as is the case of a CO₂-induced global warming, the situation has never happened in the past.

The two perspectives of the CO₂ problem mentioned earlier — event and process — are useful for a better understanding of not only the problem and its impacts but of the political decision-making responses to it. For example, perceiving the increase in CO₂ levels as an event may be valuable for creating an awareness of the potential consequences. It may also be helpful in generating public and political concern about the risks associated with a continued dependence on fossil fuels. However, if the analogy of urban air pollution as a microcosm of the CO₂ problem is accepted, then it can be seen that spectacular arguments about the CO₂ doubling will not bring about the desired public policy decisions concerning alternative energy systems.

Muddling through on long-term problems

What is needed is a decision-making approach that matches the environmental situation. While the concern with a doubling might initially create a crisis decision-making atmosphere, it would be an ineffective approach to an environmental problem which is essentially low level but cumulative, a not so spectacular problem that requires sustained interest over a long period of time.

In the past the decision-making approach to low level environmental problems such as air pollution has generally been one of muddling through. That may have been adequate in dealing with some political and social problems, but it has not been an effective way to deal

with low-level cumulative environmental problems. Eventually such problems increasingly worsen until a crisis situation is perceived to prevail. Then policy making grows out of crisis management.

Crisis decision making from the outset has also been ineffective as a way to deal with low-level cumulative insults to the environment. After a brief flurry of decision-making in a crisis atmosphere, the crisis (perceived high threat, short time to act, high cost) in this case does not immediately materialise. As a result, interest in the problem dissipates and those who again attempt to generate interest receive little response from the media, the public and the politicians. Environmental degradation, however, continues to worsen steadily.

Crisis thinking may be the only way

What is needed is a way to deal with the two aspects of the CO₂ problem simultaneously. The CO₂ problem, not unlike the more localised urban air pollution problem, might be seen as an "impending" crisis and, therefore, in need of a special type of decision-making process, one that combined the positive elements of a "crisis awareness" with cautious decision-making.

An interesting question that arises is whether a democratic society can effectively deal with a CO₂-like environmental problem. Of the competing interests in such a pluralistic society, some of those interests will be winners and some will be losers. The demand of society for continually increasing amounts of energy derived from the burning of fossil fuels may be more immediate and pressing than the environmental effects of such activities. In addition, in every society where resources are scarce, there is a great competition among interest groups for attention, time, and resources to deal with their particular problem. In light of the hydra-headed crises that confront most societies (water, food, population, energy, disease, etc.) interest in any one problem (especially the low level ones) lasts only for a short time before society shifts its attention to the other problems that confront it. So it is possible that a pluralistic society may best be equipped to deal with environmental problems on a crisis basis, that is, it is best (if not only) capable of fighting brush fires as they arise.

The question then becomes one of developing a policy-making approach for a pluralistic society (including the international community), one which can sustain the interest and concern of various segments of society in a low-level but cumulative environmental problem such as the effects of the continually increasing burning of fossil fuels. □

The author is a political scientist at the National Center for Atmospheric Research at Boulder, Colorado, USA.

news and views

Longer perspectives in neutrino physics

from D. J. Miller

THE honeymoon is drawing to an end. The charms of the 'new physics'—weak neutral currents, the J/ψ , the upsilon, parton-like scattering and jets—are now part of the everyday world. At the Bergen conference 'Neutrino '79'* this change of outlook showed in a number of ways.

J.D. Bjorken (Stanford Linear Accelerator Center) summarised the situation. A year ago individual experiments, such as the Gargamelle neutrino-electron scattering result (see *News and Views* 273, 98; 274, 531; 1978) looked as if they might change our whole understanding of the field. But as such fluctuations are superseded by accumulating statistics, the 'standard model' has shown great elasticity. The same picture always reasserts itself. The strongly interacting particles (hadrons) are explained very economically in terms of fractionally charged 'coloured' quarks, fundamental particles at the same level as the leptons (electron, muon, tau and the neutrinos), but confined in threes or in quark-antiquark pairs. The electromagnetic and weak couplings of quarks and leptons have settled into a simple pattern. Last year's questions were about the neutral weak currents. This year three theorists stood up one after another (J.J. Sakurai, University of California, Los Angeles, P. Langacker, University of Pennsylvania and M. Roos from Helsinki) and, with a variety of starting assumptions—discarding part of the data, varying the theoretical model, fitting 10 parameters, 7, 5, 2 or 1 parameter—they all came back to the same values for the couplings. There is still a lot of checking to be done, but many of the experimental reports on neutrino results from accelerators are beginning to sound routine.

All theoretical discussions of the immediate experimental picture are dominated by the two gauge theories which together now constitute the standard model; QCD (quantum chromodynamics, see Marciano and Pagels *Nature* 279, 479; 1979) for the strong interaction with quark-binding, and $SU(2) \times U(1)$ (the 'Glashow-Weinberg-Salam model', see *News and*

Views 279, 675; 1979) for the weak and electromagnetic interactions. Neither of them is proved to be correct but each offers an economical and comprehensive explanation of a large group of phenomena. Neutrino results which test the predictions of QCD were reported by a number of groups and it is clear that the breakdown of the simplest quark-parton model of neutrino or muon scattering is consistent with QCD. A few results, notably some new muon scattering data from Fermilab experiments, did not quite fit in—but people were inclined to wait and see whether such disagreements would just go away, as so many others have done.

Two working gauge theories would seem to imply an underlying unified gauge theory. 'Grand unified models' are being put forward with increasing confidence—the simplest being the $SU(5)$ theory due to Georgi and Glashow. The reaction to such models last year was lukewarm. How could we test a unified theory when the forces it describes (strong, weak and electromagnetic) only become equal at an energy of around 10^{15} GeV? Now, however, it has been realised that an important experimental test lies within easy reach. The credit for pointing this out belongs, among others, to Cecilia Jarlskog of Bergen—chief organiser of the conference and a uniquely good-humoured review-speaker. In her talk she reminded us that if the 'grand unification mass' is 5×10^{14} GeV then the proton lifetime should be between 10^{30} and 10^{33} years (see *News and Views* 278, 687; 1979). Experimental limits on this lifetime are already more than 10^{29} years—a by-product of cosmic-ray neutrino experiments. The most interesting experimental session of the conference to me was the one which contained no new results at all but in which L.R. Sulak (Harvard University) and E.C. Fowler (Purdue University) talked about plans for experiments to push the limit on the proton lifetime to beyond 10^{33} , perhaps even to 10^{35} years. Sulak presented a scheme which is likely to be working within a year. It involved no new technology—just a tank of water $20 \text{ m} \times 20 \text{ m} \times 20 \text{ m}$, with a grid of photomultipliers over all of its faces to detect Cerenkov light from the electromagnetic showers which would be

produced by such decays as $\text{proton} \rightarrow \text{positron} + \pi^0$.

The implications of the new physics, in particular those of the neutral weak currents, are now being worked out in other fields than high energy particles. An important series of experiments is taking place on the quantum-optical effects of parity violating couplings. E.D. Commins (University of California, Berkeley) reported that the Novosibirsk result on parity-violating optical rotation in bismuth vapour is looking increasingly secure and is agreeing with the $SU(2) \times U(1)$ model predictions. He presented his own preliminary but positive results on dichroism in thallium and he suggested that the null results seen by some groups working with bismuth may have been due to such subtleties as variable side-bands in their laser beams. The present motivation for all such experiments is still to check the weak neutral current couplings, but I wonder if this is not likely to become a 'spin-off' field from particle physics, comparable to the 'spin-offs' of the Mossbauer effect and nuclear magnetic resonance from nuclear physics. Perhaps in a dozen years optical parity-violating effects will be another tool in atomic physics or in chemistry; the apparatus is not expensive and laser technology is advancing rapidly. Low energy nuclear physicists are also now joining in the search for neutral-current effects and some positive results were reported.

But astrophysicists provided the widest horizons. There is still a problem in the observation of neutrinos from the Sun. M.S. Davis's group from the Brookhaven Laboratory see about one-half or one-third of the predicted rate of transitions $^{37}\text{Cl} + \nu_e \rightarrow ^{37}\text{Ar} + e^-$. This reaction is only sensitive to higher energy neutrinos which come mostly from a subsidiary branch of the solar nucleosynthesis cycle. For years there has been a debate as to whether the solar models, the nuclear transition rates or the properties of neutrinos were to blame for the anomaly. Now a new collaboration has been formed—reported by Kirsten from Munich and including Brookhaven,

D. J. Miller is in the Department of Physics, University College, London.

* 'Neutrino '79' was held in Bergen on 18–22 June and organised by the University of Bergen and NORDITA.

Pennsylvania and the Weizmann Institute, which will look for the more copious lower energy neutrinos that must be produced by the basic solar hydrogen-burning process $p + p \rightarrow d + e^+ + \nu_e$. They will collect 50 tons of gallium over the next few years and search for the expected production of one atom per day of radioactive germanium by way of the process $^{71}\text{Ga} + \nu_e \rightarrow ^{71}\text{Ge} + e^-$.

Other astrophysicists reported on the apparent regress of the theory of stellar collapse. This has followed from a serious treatment of the equation of state of highly condensed matter. It had been thought that neutrinos could provide the mechanism for transporting energy from the core of the supernova to the exploding envelope, thus allowing the collapsed core to get rid of its binding energy and form a neutron star. But present values for neutral and charged current cross-sections require that during the crucial last second of collapse the neutrinos are trapped in the core and form yet another degenerate Fermi gas, in equilibrium with neutrons, protons, electrons and nuclei. As a consequence of this, a number of speakers reported, it is not now possible to show how a supernova envelope can be formed by the collapse of a massive star. New ideas and calculations are clearly needed, but the theorists are hampered by a poverty of experimental constraints. Supernovae are seen only about once per century in our Galaxy and no one has yet observed the expected 'neutrino flash' that should go with the visible flash. Such an observation would be extremely useful.

The biggest accelerator of them all, the primeval big bang, provides the most speculative field of application for the new physics. Cosmologists, D. Schramm (University of Chicago) told us, have embraced the Grand Unified models as a potential answer to the question of why the universe seems asymmetric in baryon number. Why is all stable matter apparently not matched by equal amounts of antimatter? If the massive (10^{15} GeV/ c^2) particles exist which would mediate proton decay they must have been produced copiously in the very earliest stages of big bang. Their decays could be affected by substantial CP symmetry violation, leading, as in K^0 processes, to a time asymmetry in the production of positive and negative particles. In thermodynamic equilibrium, such effects would be cancelled out, but in an expanding and cooling system an overall baryon excess over antibaryons might well be produced. Variations on this theory have been suggested, involving such contracting mechanisms as a 'foam' of mini black holes or supercluster-sized zones of baryon and anti-baryon matter caused by local differences in spontaneous symmetry-breaking.

In the final sessions Glashow (Harvard University) gave his customary shopping list of work still to be done. Despite the widening influence of recent results, there

are many gaps in our understanding. In particular, why do we have the same sort of family of quarks and leptons three times over? Why is each generation so much heavier than the last? How many more generations are there? (If more generations mean more massless neutrinos then the cosmologists cannot account for the amount of helium in the universe). Accelerator and storage-ring physics may spend the next few years tidying up at the present level of understanding but we

should be building new machines at the same time. The CERN pp collider, the Isabelle pp collider at Brookhaven, the suggested European LEP e^+e^- machine and an electron-proton collider are all needed if we are to penetrate into the world of gauge bosons which promises to exhibit a rich variety of phenomena, as well as leading us closer to a unified model. Though, as Glashow reminded us, gravity should also be included before we have the truly 'Grand' unified theory. □

Recognition, restriction and immunity

by Miranda Robertson

THE immune response of animals is controlled by two sets of genes: the immunoglobulin variable region genes (V genes) which determine the responses of lymphocytes to antigens; and the major histocompatibility complex genes (MHC genes), which determine their responses to one another. It is now clear that while B cells see antigen *per se*, by means of surface immunoglobulin coded by V genes, T cells see antigen only in association with the appropriate syngeneic MHC determinant, and it is this combination that induces the helper, suppressor or killer response. But whereas the molecular basis of antigen recognition by B cells is well understood, that of T cells has remained elusive. The attempt to define at the molecular level both what T cells see with, and exactly what they see, was therefore one of the main preoccupations of participants at a recent meeting on recognition and function in T and B lymphocytes.*

One crucial question is the contribution of the V genes to T cell recognition. Eichmann and Rajewski, and Binz and Wigzell have provided strong circumstantial evidence for the expression of V genes in T cell receptors with the demonstration that both helper and suppressor responses to specific antigens can be induced by antibodies against V-gene coded antigenic determinants (idiotypes)^{1,2}. This is important not only for the light it sheds on T cell receptors, but also because it implies in principle that T cells as well as B cells may be subject to anti-idiotypic control as proposed in Jerne's network theory (reports of whose demise³ are greatly exaggerated).

More recently, D. Givol (Weizmann Institute) has raised antisera in rabbits that react with framework determinants of V_H chains of mouse immunoglobulin⁴. These antisera will prevent the binding of antigen to some T cells and inhibit T cell help *in vitro*, and have recently been exploited by T. Tada (Tokyo University) in the

characterisation of an antigen-specific suppressor factor isolated from a T cell hybridoma⁵.

Soluble factors derived from T cells have been the focus of much recent investigation on T cell functions. It is not yet clear how they are related to the T cell receptor: they may, by analogy with surface and secreted immunoglobulin on B cells, be a secreted version of the membrane receptor; or they may be the receptors themselves, shed into the culture medium *in vitro* but with no normal function as a soluble molecule *in vivo*.

The data from Tada's laboratory strongly suggest that the T cell receptor and the soluble factor are very closely related, and that they are determined both by immunoglobulin V genes and by genes in the I region of the MHC. The basis for this conclusion is that antisera against 'I-region-associated' (Ia) antigens, as well as anti-V_H antiserum and specific antigen can all bind both to suppressor T cells and to the antigen-specific suppressor factor. A further extremely important point is that the Ia antibodies are specific for the mouse strain that donated the suppressor cell, and the factor acts only on cells from mice with the same I-J antigens. This so-called restriction by MHC antigens is characteristic of T cell interactions and it is therefore important to know whether it applies to soluble factors.

An important issue which hybridomas should help to resolve is that of the part played by antigen nonspecific factors in lymphocyte interactions. Although there have been many reports of such factors, some feel that such nonspecificity in a system whose principal characteristic is exquisite specificity must make their functional relevance suspect. Others, including Tada, have incorporated them into hypothetical nets in which both specific and nonspecific factors play a part, acting on different subsets of cells. Tada's laboratory has developed a number of hybridomas, some producing specific and some nonspecific suppressor factors. The fact that the nonspecific factors are

The ICN-UCLA Symposium on T and B Lymphocytes — Recognition and Function, was held at Keystone, Colorado, on 26-30 March.

monoclonal eliminates the possibility that their nonspecificity is apparent rather than real, and that the cells are really producing a very large number of different specific factors.

A similar collection of hybridomas producing helper factors would be a great asset. The biochemical elucidation of helper factors lags well behind that of suppressors. There is now evidence for idiotype-bearing, antigen-specific factors which replace T cell help in the production of antibodies against phosphorylcholine (M. Feldman, University College London) and (T,G)-A--L (ref. 6) but the relationship of such specific factors to nonspecific helper factors remains unclear. J.D. Watson (University of California at Irvine), for example, reported that T cell growth factor(s) from supernatant of spleen cells treated with concanavalin A would act both as a nonspecific growth factor and as a nonspecific helper factor in antibody production or the generation of cytotoxic T cells; but if microcultures of ten or so cells are expanded with the use of the growth factor, their supernatants are active only on some cells. This experiment, which is similar to that of Waldmann and Lefkowitz⁷ suggests that 'nonspecific' growth factors from concanavalin A supernatants may be mixtures of antigen or idiotype-specific factors; it remains to be seen whether such mixtures include a genuinely nonspecific T cell growth factor.

MHC restriction, which limits T cell recognition of viral antigens to infected cells bearing the same MHC antigens, can be seen as a special case of antigen specificity applied to T cells. It has now become clear that, like any other kind of antigen recognition, MHC restricted recognition is subject to cross-reaction. Cross-reactions have been demonstrated in experiments in which allogeneic virus-infected cells are killed by cytotoxic cells depleted of alloreactivity by passage through the appropriate mouse strain (J.R. Bennink, University of Philadelphia). The cross-reaction may not be reciprocal: for example, b haplotype cells may kill infected k haplotype targets while k cells will not kill infected b targets⁸.

The converse can also occur: certain combinations of haplotype and virus can fail to elicit a cytotoxic response even when MHC restriction requirements are met. Much of the interest in such failures of recognition is of course due to their possible bearing on susceptibility to disease in man. A McMichael (Oxford University) reported that human lymphocytes of certain haplotypes will not show MHC-restricted killing in association with influenza virus antigens, whereas others do. Similarly, S. Shaw (US National Institute of Health) reported that some HLA haplotypes elicit much stronger anti-influenza cytotoxic responses than others.

These data lend support to the proposal that the immense polymorphism of the MHC antigens is due to the massive selective advantage of heterozygosity: if the combination of one parent's MHC antigen and virus does not elicit a response, the other parent's may (indeed this has been shown⁹).

However, this begs the question of why antigens that fail to elicit an immune response should survive in the population. The answer may lie in the mechanism of selective recognition. It is now clear that some combinations of MHC antigen and virus cross-react with alloantigens¹⁰; it is thus highly likely that the combination of some viral antigens with self antigens will cross-react with other self antigens. Since any cell reacting strongly against self must be eliminated, it follows that cells reacting against certain combinations of self and virus will sometimes be eliminated too.

While it is clear that to avoid autoimmune reactions T cells with high-affinity receptors for self must be eliminated, it has been suggested that receptors with weak affinity for self antigens are the basis for MHC restricted recognition¹¹. Tentative evidence in support of this theory was presented by H. Wigzell (Uppsala University), who has used an anti-idiotypic serum raised against immune T cells specific for an allogeneic antigen to extract the T cell receptor for that antigen. He then labelled internally the allogeneic antigen, the self antigen of the immune rat and a third-party antigen, and ran them on Sepharose columns to which the receptor was bound. Three labelled polypeptides from the third-party control passed straight through the column; two chains from the cells of the immunising strain were stopped; and one chain from the self strain was retarded.

Wigzell interprets this as evidence not only for weak binding of alloreactive receptors to self, but also as evidence that different polypeptides are involved in the recognition of self MHC and allo- or viral determinants. It is still a matter of dispute whether T cell receptors 'see' virus and MHC antigen as a complex, with only one receptor for both ('altered self'), or whether they see the two antigens with different receptors ('dual recognition')¹². Attempts to resolve this question at a cellular level are all open to doubt: it is probable that only biochemistry can produce the answer, and Wigzell's experiment may be a first step. □

¹Eichmann, K. *Eur. J. Immun.* **5**, 511-517 (1975).

²Eichmann, K. & Rajewski, K. *Eur. J. Immun.* **5**, 661-666 (1975).

³Langman, R. *Nature* **277**, 517 (1979).

⁴Ben-Nevia et al. *Eur. J. Immun.* **8**, 797 (1978).

⁵Taniguchi, M. et al. *Nature* (in the press).

⁶Mozer, E. & Haimovich, J. *Nature* **278**, 56-57 (1979).

⁷Lefkowitz, I. & Waldmann, H. *Immunology* **32**, 915 (1977).

⁸Doherty, P. & Bennink, J.R. *J. exp. Med.* **149**, 150-157 (1979).

⁹Shaw, S. & Biddison, W.E. *J. exp. Med.* **149**, 565-575 (1979).

¹⁰Finberg, R., Burakoff, S.J., Cantor, H. & Benacerraf, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5145-5149 (1978).

¹¹Janeway, C.A., Wigzell, H. & Binz, H. *Scand. J. Immun.* **5**, 993-1001 (1976).

Ecology of photosynthesis

from Peter D. Moore

Resource partitioning in plants is less well documented than among animals and this is not surprising, for the resources on which plants depend are fewer and adaptations which could be referred to as partitioning are generally more subtle than is the case with animals. This is probably why plant ecophysiologicalists have eagerly grasped the opportunities presented by the gradual exposition of the extent of the C_4 and crassulacean acid metabolism (CAM) photosynthetic adaptations among plants. It is well established that both of these modifications to the basic C_3 Calvin photosynthetic system are associated with certain ecologically selective advantages in conditions of high temperature, high light intensity and water stress. This has led to many investigations into the biogeographical significance of the adaptations, mainly by mapping their distributions on a continental scale (see *Nature* **272**, 400; 1978). Information is now accumulating concerning the ecological value of such adaptations in the local composition of plant communities.

To search for correlations between the success of one particular type of photosynthetic system and an environmental gradient, such as aridity or elevation, is not new. In 1974, Mooney, Troughton and Berry (*Carnegie Inst. Ybk.* **73**, 793; 1974) reported a survey they had conducted along the coastal strips of California and Chile, where the climate shifts from a Mediterranean to a desert type with decreasing latitude. They found in both locations that increasing aridity was associated with an increasing proportion of CAM plants in the flora. Among C_3 plants, there was a change from evergreen to drought-deciduous forms, but C_4 plants were virtually absent from all stations. They explain this as a consequence of the low temperatures prevailing at the times of water availability.

More recent work in the cold winter deserts of Utah by Caldwell, White, Moore and Camp (*Oecologia, Berl.* **29**, 275; 1977) has also shown that such C_4 species as *Atriplex confertifolia* fail to achieve many of the advantages normally associated with the C_4 system when they commence growth in the cool of spring. They may, however, still match the growth of their C_3 competitors and gain the advantage of a longer growing season as summer aridity sets in. It is difficult, therefore, to explain the lack of C_4 species in the Californian and Chilean coastal deserts.

Eickmeier has now published the results of a similar environmental gradient study from Texas (*Photosynthetica* **12**, 290;

Peter D. Moore is a Senior Lecturer in the Department of Plant Sciences, King's College, London.

1978). Of 88 non-herbaceous species found along a 1,300-m elevation gradient, 22 were CAM, three were C_4 and the remaining 63 were C_3 . Ten samples were taken at 150-m altitude intervals; density, frequency and area cover values were used to determine an importance value for each species at each elevation. The main peak in the importance of CAM species was found to occur at low altitude, that of C_3 species at higher altitudes and the C_4 peak was intermediate between the other two. Eickmeier considers this sequence to reflect both the adaptational significance of the physiological variations and also the respective responses to biotic, competitive factors. For example, the CAM plants are most suited to the low altitude, arid environment, where competitive interplay with other plant species may be regarded as minimal (but see *Nature op. cit.*). C_3 plants, he claims are evolutionarily attuned to competitive situations with low abiotic stress; C_4 plants can cope with both, having potentially high productivity and the

capacity to withstand water stress.

The greater competitive success of C_3 species in higher altitude, more mesic conditions in the tropics has now been demonstrated by Tieszen, Senyimba, Imbamba and Troughton (*Oecologia, Berl.* 37, 337; 1979) along an altitudinal and moisture gradient in Kenya. They have found that nearly all the grass species of the low altitude grasslands in Kenya are of a C_3 photosynthetic type. Some grass tribes, such as Paniceae and Andropogoneae (both exclusively C_4) were most frequently encountered at intermediate altitudes. The exclusively C_3 tribes (such as Festuceae, Aveneae and Agrostideae) were found only at high altitude. No C_3 species were found below an altitude of 2,000 m and no C_4 species above 3,000 m.

Evidently the complex adaptational significance of C_3 , C_4 and CAM photosynthetic systems is profoundly influential on the floristic composition of plant communities both on a biogeographic and an ecological scale.

Viral pathogenesis and virulence

from D.A.J. Tyrrell

It is easy to say "The influenza virus causes influenza", while having really no scientific explanation of many of the signs and symptoms of the disease, for instance the headache or the depressed white blood cell count. This meeting* would have jolted anyone out of such a complacent view and also gave some good examples of how patient, thoughtful research is building a more complete picture of what happens in some virus infections — how the virus enters and spreads through the body, enters cells, disturbs their function, interacts with the host defence mechanisms (such as interferon or humoral or cellular immunity) and disturbs normal body functions to produce what we recognise as a clinical disease.

What seem in the laboratory to be very similar viruses may nevertheless produce either different diseases or similar diseases by different methods. There was for instance a session on coronaviruses, which all look the same and are similar in biochemical composition. Yet the different coronaviruses are specific to particular species — mouse coronaviruses never seem to affect man or *vice versa*, for example — and sometimes even to a very limited variety of cells in one species. P.A. Bachmann (World Health Organization, Munich) described how the coronavirus of transmissible gastroenteritis of pigs (TGE) affects only the differentiated enterocytes of the villi of the small intestine which may almost disappear. The crypt cells,

however, survive and migrate up to cover the villi again, but while they are doing this and differentiating into mature enterocytes, there are profound physiological disturbances — lack of absorption, excess secretion, lack of digestive function (for instance the splitting of lactose by disaccharidase) — leading to marked diarrhoea, which can kill a piglet from dehydration unless it is treated. On the other hand, another pig coronavirus causes vomiting and wasting disease. M. Pensaert and K.L. Andries (State University of Ghent) described how this virus affects the upper respiratory tract and lung, the tonsil and the jejunum; it affects the cells of the plexus of Auerbach and Meissner in the gut and seems to spread from them. It spreads through the nervous system to the brain stem and later to the spinal cord and the rest of the brain. All this was established in great detail by painstaking serial virus assay and immunofluorescent microscopy of experimentally infected animals. This suggests how the disease develops — probably the virus spreads from the respiratory tract and gut along peripheral nerves to the brain stem where the lesions give rise to vomiting, while the damage to neurones in the plexuses prevents normal peristalsis, leading to wasting. Furthermore, knowing that the virus invades the nervous system widely, it is not surprising that a virtually identical virus has been recovered from pigs diagnosed as having encephalitis, and so was named haemagglutinating encephalitis virus.

Explanations of these important and mysterious tropisms of viruses must lie at

the molecular level, in the functioning of the peptides and nucleic acids of which viruses are made. B. Fields (Harvard Medical School) gave an example of what can be done, when he described his study of two pairs of genes, represented by the RNAs S1 and M2 of reoviruses types 1 and 3. He studied parent viruses and a collection of recombinants (or reassortants) and showed that S1 nucleic acid specifies the haemagglutinin protein of the virus and that this is responsible for the ability to agglutinate red cells and for serological specificity. It is also responsible for cell tropism in the animal, so that the type 3 virus causes a necrotising encephalitis due to a specific attack on the neurones. It also determines the extent to which the virus inhibits DNA synthesis when absorbed to cells. M2 RNA on the other hand, which is abundant on the surface of the virus particle determines whether the virus resists inactivation by chymotrypsin. M2 RNA from type 3 makes the virus resistant to host enzymes and hence able to infect when given by mouth. Discussion of this work showed that the next step was to determine at which site on the surface of the neurone the virus attached. This is probably a part of the chemical matrix which has some other physiological function characteristic of neurones. Reassortants were also made which were infectious when given by mouth, like type 1, and caused central nervous system disease, like type 3, and so, by the oral route, were more pathogenic than either parent. This type of manipulation confirms earlier conclusions about the functions of the genes.

It is very difficult to evaluate the way in which the host defences interact with the virus, sometimes preventing its multiplication, sometimes causing immunological damage to infected tissue. For instance, P.A. Neighbour and B.R. Bloom (Albert Einstein College of Medicine, New York) described some of their work which started when they analysed why the lymphocytes of patients with multiple sclerosis show an unusual reactivity to measles virus; they tracked down the fault to a defect in the ability of their lymphocytes to produce type 1 interferon (Neighbour & Bloom *Proc. natn. Acad. Sci. U.S.A.* 76, 476; 1979).

The success of the meeting was obvious, and there were many other interesting papers besides those mentioned here, including whole sessions on herpes viruses and rhabdoviruses. There were clear signs that pathogenesis not only holds scientific interest, but may be a source of practically useful knowledge: vaccines and antiviral treatments do not always work — sometimes because we guess too much and know too little about the process by which viruses produce disease. □

D.A.J. Tyrrell is in the Division of Communicable Diseases, Clinical Research Centre, Harrow.

*A symposium on Mechanisms of Viral Pathogenesis and Virulence was held in Munich on 6 and 7 June, 1979 and organised by P.A. Bachmann.

Seeing is believing in lineages

from a Correspondent

At last a directly observed and accurate cell lineage is available for the development of one organism, the nematode *Caenorhabditis*. As reported at a recent meeting*, the lineage runs from the fertilised zygote to the completion of embryogenesis (550 cells), and then there is a gap of one or two cell divisions, which have not yet been studied, before the lineage runs directly through to adult structures. This description is now up for analysis and G.von Ehrenstein (Max-Planck Institute for Experimental Medicine, Göttingen) has isolated at least seven temperature-sensitive mutants whose maternal gene expression is necessary and sufficient for normal embryogenesis, and which eventually block development at the 'lima bean' stage. Three of these produce abnormal guts without obviously disturbing the rest of the lineage; in the gut lineage one changes the direction of a single division while the other two change the relative timing of division and migration of gut primordial cells to the interior. It is, however, unlikely that the regularity of the lineage merely depends on control of cell division rate and cell division direction for there are even more extensive cell migrations across four or five cell diameters in other parts of the embryonic lineage. Certainly there is regulation in the post-embryonic lineage for, following laser ablation of single cells, their lineage and adult structures may be replaced by neighbouring cells within an 'equivalence group' (J. Kimble, MRC Laboratory of Molecular Biology, Cambridge). It remained unclear whether the replacing cell had to repeat the lineage of the ablated cell in order to replace its adult structures because within an equivalence group it always did so. (This and related work with *Caenorhabditis* was mentioned in *News and Views*, 279, 758; 1979.)

The masterplans in 'mosaic' eggs were certainly under attack even though there are nice sets of computer simulations which could manufacture such plans (E. Parisi, Laboratory of Molecular Embryology, Naples). P. Guerrier (Biological Station, Roscoff) reviewed his own and van den Biggelaar's beautiful manipulation experiments on the 'mosaic eggs' of annelids and molluscs and made it difficult to see how anybody could still believe in mosaicism in these embryos (for example, Biggelaar & Guerrier, *Devl Biol.* 68, 462; 1978). The *Drosophila* egg must at least be allowed its polarities, and to go with 'bicaudal', there is now a maternal effect mutant, dorsal (*dl*), which leaves the anterior-posterior polarity intact but which lacks lateral and ventral structures

(C. Nüsslein-Volhard, EMBL, Heidelberg). It may be 'all dorsal' for no dorsal midline can be detected, and in order to show formally that the laser ablation fate map was altered by this dorsal hypodermis embryo it was necessary to put the mutation on a background which reduced its penetrance so that some ventral structures were formed. The polar tendencies of the *Drosophila* egg must arise during oogenesis and E. Weischaus (EMBL, Heidelberg) has now made use of another maternal effect mutant, *fs(1)K10*, in a neat mosaic analysis of stem cell behaviour during oogenesis. This mutation is expressed as a chorion defect (allowing immediate recognition of affected eggs) but its expression is restricted to the germ line; consequently recombination induced by X-irradiation during the formation of eggs can be used to follow the behaviour of stem cells which are egg precursors. From blastoderm to pupa, the stem cells increase to about 100 and this level is maintained through egg production. This technique could be exceptionally useful for analysing the effect of characters expressed in the nurse-cell oocyte clusters on subsequent development.

Even those working on haemopoietic stem cells could now see lineage diversification with their eyes and could provide direct evidence for multipotential stem cells. G.R. Johnson (Walter and Eliza Hall Institute, Melbourne) has now picked single cells, waited for them to divide twice in isolated liquid culture, separated the four cells, found that each could form colonies which contained a variety of cell types, and thus showed that one cell could form erythrocytes, macrophages and granulocytes. The implication is that cell lineages can separate without the special cell contacts and microenvironments which are supposed to pick out and diversify stem cells; nobody could decide if the behaviour of these clones was natural or perverted. The environmental lobby made a comeback with a study of the progressive education of T lymphocytes to attack virus-infected cells. It has now been shown that the major histocompatibility complex (MHC) type of both the thymus and the 'post-thymic environment' educates cytotoxic T cells to kill virus-infected cells with similar MHC (R.M. Zinkernagel, Scripps Clinic, La Jolla).

If lymphocytes can be educated, then nerve states can be altered and stabilised. The evidence that individual post-mitotic neurones can functionally convert from adrenergic to cholinergic cells in the presence of non-neuronal cells is now convincing and P.H. Patterson (Harvard University) went on to show that this change could be prevented by depolarising the neurone. This change could be seen directly in cell culture; could it be as simple as that in the adrenergic-cholinergic switch shown by neural crest cells transplanted to different levels of the axis of developing chick-quail chimaeras? Certainly these

transplants illustrate beautifully the importance of the migratory path and final location in picking out the phenotype of neural crest cells (N. Le Douarin, Embryology Institute, Nogent-sur-Marne). Depolarising a neurone in culture may mimic normal neurone activity, and chronic stimulation of the spinal cord during embryogenesis may do the same. If it does, then the stabilisation of nerve synapses on muscle during development may be shown to depend on nerve activity. The system is the posterior latissimus dorsi muscle of the chick which early in development has several synapses per fibre and of these only one remains in the adult; stimulation of the spinal cord during embryogenesis maintains multiple connections (P. Changeux, Pasteur Institute, Paris). Is this the selective stabilisation theory of nerve connections in action? □

Mutants of the *lac* repressor

from Alan D. B. Malcolm

As part of their study of the binding of the *lac* repressor protein of *Escherichia coli* to the *lac* operator DNA, Miller and colleagues set out several years ago to produce mutants of the repressor protein by an ingenious technique, and the culmination of their work appears in three recent papers (*J. molec. Biol.* 130, 191, 223 & 249; 1979). Miller's approach has been to use chemical mutagens to produce nonsense mutations (where a coding triplet is replaced by one of the three 'stop' signals, UAG, UAA or UGA) in a suppressor strain which possesses a tRNA able to substitute an amino acid in place of the stop. So far it is possible to insert Ser, Gln, Tyr, Lys or Leu at UAG; Gln, Tyr or Lys at UAA and Trp at UGA. This method has the advantage that it is easy to screen for proteins with minimal functional defects (this is difficult to achieve by conventional mutagenic techniques) and also that some of the mutations introduced would be expected to require two base changes. It also has some limitations. Mutations can be introduced only at positions where the wild type can be mutated into one of the three 'stop' signals. It is known that not all sites are equally susceptible to mutagenesis and the distribution of these 'hot spots' has been the subject of a previous paper (Farabaugh *et al. J. molec. Biol.* 126, 847; 1978). The amino acids that may be introduced are limited by the availability of suppressor strains, although there seems to be no reason why the number of these could not be increased in the future. Miller *et al.* have collected more than 300 mutants at 90 out of the 360 positions in the *lac* repressor and these can be altered in various ways: inducer binding, ability to form tetramers and binding to DNA (either operator or

*The meeting on Cell Lineage, Stem Cells and Cell Determination was held in Seillac on 20-24 May 1979. The proceedings will be published by Elsevier-North Holland.

nonspecific) can all be altered. It has been known for some time that it is the amino end of the protein that is involved in DNA binding (for example, *Nature*, **268**, 196; 1977) and it is therefore not surprising that mutations in this region frequently abolish DNA binding (*J. molec. Biol.* **130**, 191; 1979), whereas mutations at the carboxyl end are often without effect. Indeed a chimaera between the *lac* repressor and β -galactosidase produced by a deletion mutant of *E. coli* (Brake *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 4824; 1978) which is without the last four residues at the C terminus but has instead nearly 10^5 daltons of β -galactosidase, retains all the various properties of the repressor. It is therefore interesting that when the Tyr 273 is replaced with Gln, the tight operator binding of the repressor disappears. Tyrosine residues have already been implicated in specific protein-nucleic acid recognition both as a result of model studies (Wehling *et al. Nucleic Acid. Res.* **4**, 513; 1977; Brun *et al. Biochemistry* **14**, 558; 1975) and also from work on gene 5 protein from bacteriophage fd (Coleman & Armitage *Biochemistry*, **17**, 5038; 1978) on gene 32 protein from bacteriophage T4 (Anderson & Coleman *Biochemistry*, **14**, 5485; 1976), as well as in the N-terminal region of the *lac* repressor (Alexander *et al. Biochim. biophys. Acta* **493**, 367; 1977).

Miller and Schmeissner have also analysed many missense mutations in which the use of chemical mutagens has resulted simply in the substitution of one amino acid for another (*J. molec. Biol.* **130**, 223, 1979). As expected, the concentration of i^- mutants (where there is no active repressor, so that the *lac* enzymes are constitutive) is particularly high among the first 60 amino acids. The i^s mutants are those in which the *lac* enzymes are repressed even in the presence of inducer and none of these occurs in the early part of the gene. This is exactly as expected because it is known that the binding site for IPTG (a synthetic inducer) depends on residues 60-360 and not on 1-59 (for example, *Nature* **268**, 196; 1977).

When the results of the suppressed nonsense mutations are taken together with those for the missense mutations (*J. molec. Biol.* **130**, 249; 1979) some striking regularities appear. For example, the positions where i_s mutants occur in clusters about 26 amino acids apart (that is, around 193, 219, 245, 271 and 297) both for missense mutations and for suppressed nonsense mutations. The exact significance of this is not yet clear, but Miller speculates that the inducer binding site may resemble the antigen-binding site of immunoglobulins. This is a cavity formed by residues in hairpin turns. The hypervariable regions occur at the turns

themselves and mutations here affect the antigen-binding properties without causing alteration of the overall structure of the protein. If a similar regular structure were used to make the inducer binding site then mutations at the turns (assuming them to be about 26 amino acids apart) might alter inducer binding while leaving the overall protein structure unaltered.

All this must remain speculative until physical studies define the three-dimensional structures of the repressor and its complexes with inducer and operator. Nuclear magnetic resonance (NMR) is giving increasingly detailed information about such structures. Lu and his colleagues have grown *E. coli* on 3-fluorotyrosine in place of tyrosine and have isolated a fluorinated *lac* repressor (*Proc. natn. Acad. Sci. U.S.A.* **73**, 3471; 1976) and have now succeeded in assigning most of the ^{19}F peaks in an NMR spectrum to individual tyrosine residues (P. Lu, personal communication). Comparison of the proton NMR spectrum of the intact protein with the amino tails (residues 1-51) and the remaining core, shows that the tails have the same structure whether on their own or as part of the complete repressor (Buck *et al. FEBS Lett.* **96**, 335; 1978). Further, the narrowness of the linewidths suggests that the tails have a faster tumbling rate than would be expected if they were an integral part of a protein. It seems as if there must be a very flexible 'hinge' between positions 50 and 60 which allows the core and the tails almost independent motion.

The other half of the problem of course is to study the operator DNA. Caruthers and colleagues (Goeddel *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3578; 1978) have synthesised *lac* operators substituted with nucleotide analogues such as 5-bromouracil, 5-bromocytidine, uracil and hypoxanthine. These and previous binding studies (Ogata & Gilbert *Proc. natn. Acad. Sci. U.S.A.* **74**, 4973; 1977) suggest that the repressor binds in the major groove on one side of the DNA only. Narang has used his phosphotriester method to synthesise various sections of the operator (Narang *et al. Can. J. Biochem.* **55**, 1125; 1977) and has now studied amino acid binding to the central nine base pair sequence. (Patel *et al. Biochem. biophys. Res. Commun.* **85**, 1560, 1978) The free energies of binding tyrosine and glutamine to this oligonucleotide are both about 17 kJ mol $^{-1}$. Because the free energy for binding intact repressor to intact operator is only about 73 kJ mol $^{-1}$, it seems possible that only four or five amino acid-base interactions may be required to produce this free energy; it is already known from the mutation studies that nine amino acids in the tail section are essential (including Tyr 17, Gln 18 and Gln 54).

It looks as though the *lac* repressor is beginning to yield to the combined skills of geneticists, chemists and biophysicists. □

DNA gyrase and DNA unwinding

from David T. Denhardt

WHEN DNA is replicated, transcribed, recombined or otherwise manipulated within the cell the two strands of the double helix have to be at least partially unwound. The physical problems inherent in unwinding the DNA duplex have long been recognised and led John Cairns some years ago to postulate a 'swivel' in the circular chromosome of *Escherichia coli* to enable it to be unwound. The 'swivel' has now been recast as the bacterial topoisomerases — topoisomerase I, or ω protein, and topoisomerase II, or DNA gyrase — and the latter has been implicated as an essential enzyme in DNA replication in bacteria. DNA gyrase introduces negative superhelical twists into relaxed closed circular DNA molecules; how it does this, and how this property is related to its role in DNA replication are now under intensive study. The best guess at the moment is that it relieves the positive super-twisting that builds up in front of the replication fork as the two strands are unwound during replication. Negatively superhelical DNA may also be essential for both the initiation of replication and some recombinational events.

The relation between the helical turns in a DNA duplex and the superhelical twists in a superhelical molecule are illustrated diagrammatically in Fig. 1. A closed circular duplex is equivalent topologically to a linear molecule whose ends are unable to rotate, and despite appearances all three structures in the figure have the same net number (5) of twists of one strand about the other. The two negative superhelical twists in the molecule on the right are removed by unwinding two turns in the DNA duplex; if the duplex is unwound further positive superhelical twists are introduced. To understand the action of gyrase imagine the molecule in the centre to be a DNA molecule with its usual one duplex turn per roughly 10 base pairs. Gyrase binds to the DNA molecule in such a way as to introduce compensating positive and negative superhelical twists and then it breaks and reseals the DNA duplex so as to selectively remove the positive superhelical twist. Dissociation of gyrase from the DNA leaves a negatively superhelical molecule.

Evidence for the existence of DNA gyrase was first obtained by Kiyoshi Mizuuchi and Howard Nash who found that closed circular phage lambda DNA was much more efficient at undergoing an intramolecular recombination event than was nicked circular DNA (*Proc. natn. Acad. Sci. U.S.A.* **73**, 3524; 1976); a nick (a

Alan D. B. Malcolm works in the Department of Biochemistry, St Mary's Hospital Medical School, London.

D. T. Denhardt is a Professor of Biochemistry at McGill University.



Fig. 1

single-strand break) in a DNA molecule prevents it from being made superhelical. Martin Gellert, together with Mizuuchi, O'Dea and Nash at the US National Institutes of Health, subsequently succeeded in purifying an ATP-dependent activity that converted relaxed closed circular DNA into negatively superhelical DNA (*Proc. natn. Acad. Sci. U.S.A.* **73**, 3872; 1976). They showed that the antibiotics novobiocin and its more active analogue coumermycin A₁ inhibited DNA supertwisting and DNA replication in *E. coli* by their ability to inhibit DNA gyrase; mutations in the *cou* gene simultaneously conferred resistance to both these drugs and produced a gyrase that was drug-resistant *in vitro* (*Proc. natn. Acad. Sci. U.S.A.* **73**, 4474; 1976).

At the same time Nicholas Cozzarelli and colleagues at the University of Chicago were approaching the discovery of gyrase from a different direction. They had purified the protein product — Pnal — which confers sensitivity to the antibiotic nalidixic acid on *E. coli*. Because nalidixic acid inhibited DNA gyrase *in vitro* Gellert suggested to Cozzarelli that Pnal might be related to DNA gyrase, and this was soon proved to be so.

Functional gyrase is a 400,000 molecular weight tetramer composed of two pairs of identical subunits of molecular weights of about 105,000 and 95,000 (Higgins *et al.* *Proc. natn. Acad. Sci. U.S.A.* **75**, 5960; 1978). These polypeptides are coded for by the *nalA* and *cou* genes respectively; it has been proposed that the genes now be called *gyrA* and *gyrB*. Either subunit alone is inactive but they readily associate to form the native enzyme. Leroy Liu and James Wang, at Harvard University, purified the analogous enzyme from *Micrococcus luteus* and the two enzymes seem to have similar properties (*Proc. natn. Acad. Sci. U.S.A.* **75**, 2098; 1978).

What does gyrase do? Dissection of its activity has revealed several aspects of gyrase action. First, it can bind to DNA. Second, it can hydrolyse ATP in a DNA duplex-dependent reaction; this is the activity inhibited by coumermycin (Sugino & Cozzarelli, *J. biol. Chem.* in the press). Third, in the absence of ATP it removes superhelical turns from DNA by the breakage and reunion of the two DNA strands. This is inhibited by nalidixic acid. It is not yet clear whether *in vivo*, gyrase would induce a single-stranded break only, allowing rotation of one strand around the other, or whether double-stranded breaks are transiently induced. Double-strand breaks are observed when dodecyl sulphate is added to a nalidixic acid-inhibited reaction. This has allowed preferred sites for gyrase action to be identified. The

break in one of the DNA strands is separated by four base pairs from the break in the other producing a linear molecule with protruding 5' ends to which the protein seems to be covalently bound. The 3' ends in contrast are free and able to prime a limited amount of synthesis by DNA polymerase. This provides the handle needed to determine the sequence around the cleavage site. As shown in a paper in the May issue of *Cell* there seems to be little similarity in the nucleotide sequences around the sites of cleavage in the different molecules studied so far except that on at least one strand, cleavage usually occurs between a T and G (Morrison & Cozzarelli *Cell* **17**, 175; 1979).

The binding of gyrase to DNA has been studied in detail and strong evidence suggests that the DNA is wrapped around the outside of the enzyme. Liu and Wang bound nicked circular duplex DNA to gyrase in the absence of ATP and then closed the nick with polynucleotide ligase. When gyrase was removed they found that the DNA had acquired positive superhelical twists. This implied that either the DNA wraps around the protein or that when gyrase binds to DNA it induces a winding of the helix. As there is no precedent for the latter but the precedent of DNA wrapping around histones in eukaryotic DNA for the former this explanation seemed more likely. They have since shown that when DNA is in a complex with gyrase a 143 base-pair segment is protected from digestion by *Staphylococcus* nuclease. When the gyrase-DNA complex was digested instead with DNase I, electrophoresis of the protected DNA after denaturation revealed a set of six single-stranded molecules of lengths, 47, 57, 67, 77, 86 and 96 nucleotides. It seems likely that the DNA is lying largely on the surface of the

enzyme in a form inaccessible to staphylococcal nuclease but partially accessible to DNase I such that the latter enzyme can nick the duplex over a limited region at one of several sites separated by 10-base pair intervals (Liu & Wang, *Cell* **15**, 979, 1978).

Using a similar approach Peebles *et al.* (*Cold Spring Harbor Symp. quant. Biol.* **43**, in the press) also confirmed that gyrase introduced positive superhelical twists into relaxed closed circular DNA in the absence of ATP. They then went one step further. To the gyrase-DNA complex they added an ATP analogue which although it induces a conformational change in the gyrase molecule cannot support the cleavage reaction. Subsequent treatment with nalidixic acid and dodecyl sulphate to induce double-strand breaks showed that reaction with the ATP analogue had shifted the cleavage site some distance along the DNA molecule. Also, about half the positive superhelical twists introduced into the DNA on gyrase binding had disappeared.

There is now sufficient information to construct a model for gyrase action. All its elements are common to one or more of the models suggested by the several groups studying the enzyme. As proposed by Wang, on binding the DNA wraps around the enzyme. As proposed by Cozzarelli (Sugino *et al.* *Proc. natn. Acad. Sci. U.S.A.* **75**, 4838; 1978) ATP then interacts with the enzyme to induce a conformational change that results in the translocation of the DNA relative to the enzyme and formation of a positive superhelical loop of DNA. As proposed by Gellert *et al.*, (*Cold Spring Harbor Symp. quant. Biol.* **43**, in the press) a nicking-closing cycle (or breakage and reunion if a double-strand break is transiently introduced) removes the positive superhelical turn in the isolated section of DNA. ATP hydrolysis returns the enzyme to its initial conformation and leaves the DNA with a net negative superhelical twist. The end result is that the DNA has been partly unwound. The similarity of this mechanism to other ATP-coupled energy-transducing processes has been discussed by Sugino *et al.* (*Proc. natn. Acad. Sci. U.S.A.* **75**, 4838; 1978) and Peebles *et al.* (*Cold Spring Harbor Symp. quant. Biol.* **43**, in the press) and their model is illustrated in Fig. 2. E and E' are two conformations of the enzyme.

In vivo evidence of the role of gyrase in the bacterial cell is provided by the inhibition of DNA replication by the antibiotics nalidixic acid and coumermycin and their analogues. Conditional lethal mutations have also been mapped in the *nalA* gene and possibly the *cou* gene (Kreuzer *et al.* *Molec. gen. genet.* **167**, 129; 1978; Projan & Wechsler in *DNA Synthesis: Present and Future*, 387, (ed Molineux & Kohiyama, Plenum, 1978).

The requirement for gyrase may simply be to remove positive superhelical twists

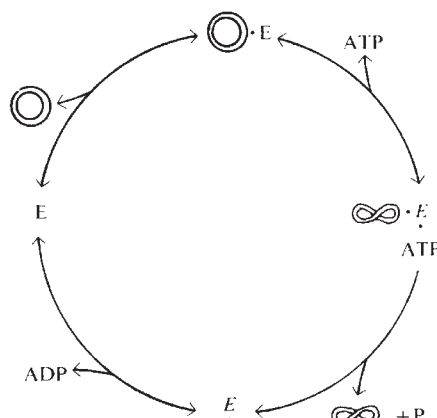


Fig. 2

induced in the DNA by the unwinding of the DNA during replication. Alternatively, or additionally, the DNA may need to be made negatively superhelical for DNA replication to be initiated. Negatively superhelical DNA is required for the initiation of replication of the bacteriophage ϕ X174 replicative form DNA *in vitro* because the ϕ X174 gene A protein, which must nick the DNA to initiate replication, does not act on relaxed DNA (Marians *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 1965; 1977).

Surprisingly gyrase seems to be required for the replication of the linear DNA of the phage T7 *in vivo* but not *in vitro* (DeWyngaert & Hinkle *J. Virol.* **29**, 529; 1979). The most likely explanation is that in the cell the DNA is collapsed into a compact structure as a result of its interaction with cellular molecules, whereas in its isolated form *in vitro* one end of the duplex is free to rotate relative to the other end and so superhelical twists cannot build up in the molecule.

McCarthy (*J. molec. biol.* **127**, 265; 1979) has discovered that the T4 DNA-delay mutants (39,52,60) require DNA gyrase for replication as they are unable to replicate in the presence of novobiocin or coumermycin in sensitive, but not drug-resistant, cells. Wild type T4 was not dependent on gyrase function, presumably because the products of one or more of the above genes form an enzyme that provides an equivalent function. From the kinetics of DNA synthesis under conditions of gyrase inhibition McCarthy deduced that the activity was required for the initiation of 'growing points' but not for the actual DNA elongation process.

On the assumption that transcription and recombination require local unwinding of the DNA it is hardly surprising that in certain cases gyrase seems to be required for both. However, the weight of the evidence suggests that gyrase is only required for transcription from some promoters (Smith *et al. Nature* **275**, 420; 1978). It is not obvious why. Harking back to the *in vitro* studies on integrative recombination in phage lambda from which the existence of gyrase was inferred it seems likely that the requirement for gyrase in recombination may be a consequence of the need for negatively superhelical DNA for this reaction, as well as for its role in DNA unwinding (Mizuuchi *et al. J. molec. Biol.* **121**, 375; 1978). One explanation for the requirement for negatively superhelical DNA is that the negative superhelical turns facilitate the uptake of pieces of single-stranded DNA to form partially triple-stranded structures. (Beattie *et al. J. molec. Biol.* **116**, 783; 1977).

It is relevant to note that those DNA metabolic processes that are coupled to gyrase action become directly dependent on the ATP content of the cell, and a high ATP content at that, since the K_m of gyrase for ATP is of the order of 300 μ M. If the cell

Oceanic heat flow, lithosphere age and thickness

from Geoff Brown

THE systematic decrease of measured heat flow with increasing age of oceanic lithosphere was well documented by the classic work of Sclater and Francheteau (*Geophys. Jl. R. astr. Soc.* **20**, 509; 1970). But in more recent years a definitive relation has been sought as proof that plate thickness increases as the square root of time. This relation follows naturally from the theory of conductive heat losses which lead to the cooling, contraction and basal freezing of oceanic lithosphere as it moves away from the zone of spreading.

A thermal model relating plate thickness, age (<120 Ma) and heat flow was summarised by Parsons and Sclater (*J. geophys. Res.* **82**, 803; 1977) who cited evidence from the North Pacific that q varies as $t^{-1/2}$ (q = heat flow, t = age). However, their data were limited to crust of age <20 Ma and >70 Ma because the reliability of heat flow and/or age control was inadequate for other times and places. In the main, uncertainties arose from the highly scattered heat flow values caused by hydrothermal disturbance in layer 1. Sclater and Crowe (*J. geophys. Res.* **84**, 1593; 1979) have taken an important step forward by measuring heat flow at 17 stations close to magnetic anomaly 13 (age = 35 Ma) on the Reykjanes ridge. By using piston cores and measuring thermal gradients across intervals down to 12 m in sediment they were able to isolate non-linear gradients above 5 m (sediments affected by recent changes in ocean bottom water temperature) from

'remarkably linear' gradients below 7.5 m. In the former case, anomalously high, non-equilibrium temperature gradients were attributed to recent changes in seawater temperature due to changes in the velocity and temperature of Norwegian Sea overflow water. The stable gradients below 7.5 m ranged from 69.9 to 102.2 mW m⁻², giving a mean heat flow for anomaly 13 of 82.9 ± 10.1 mW m⁻². The variations in these data are real in that they exceed the precision of measurement and an excellent inverse correlation, local to the survey area, was found between heat flow and depth of sediment overlying a seismic basement ridge. This suggested that either heat flow is refracted because of the contrasting thermal conductivity between layer 1 and layer 2 (that is, due to the insulating effect of the sediments) or that the interface between the layers is isothermal due to hydrothermal convection in the deeper layer: both explanations can account for the observed local variations in heat flow.

Sclater and Crowe have shown that reliable conductive heat flow data for oceanic lithosphere can be obtained by detailed and critical analysis of down-core temperature variations. Moreover, their results from the Reykjanes ridge fall on the $q/t^{-1/2}$ graph for Pacific data of other ages and provide further support for the thermal model of lithospheric thickening in proportion to the square root of age.

Geoff Brown is in the Department of Earth Sciences at the Open University.

cannot generate adequate supplies of energy the negative supertwisting of the DNA will rapidly disappear. With its disappearance those reactions which depend on having a negatively superhelical DNA substrate will not take place. The advantage of this as a control point in the prokaryotic cell is self-evident. Negative supertwisting, or its absence, is a signal transmitted throughout the entire DNA molecule, or at least that section of it that behaves as a topological unit.

So far gyrase has not been found in eukaryotic cells and there is no reason to expect it. The eukaryotic nicking-closing enzyme is sufficient to remove the positive superhelical twists induced in the DNA during replication. the superhelical twists present, for example, in purified SV40 DNA are accounted for by the number of nucleosomes in the DNA. (Each turn of the DNA around a nucleosome is approximately equivalent to one superhelical twist). Consequently eukaryotic DNA is not under an unwinding

stress resulting from the presence of negative superhelical twists and this could be a partial explanation for its being replicated more slowly than prokaryotic DNA. □



A hundred years ago

Glow-worms

Shelley sings of a "glow-worm golden in a dell of dew," but last night, at 10 o'clock, while travelling on a bridle path among the bleak lonely mountains of Tynron, Dumfriesshire, bearing up against a high wind with cold rain, I espied three glow-worms shining among the grass and ferns. I had seen them in the same locality before, but scarcely expected to have noticed them in such ungenial weather when summer has with us scarcely yet begun. J.S.

July 8

From *Nature* **20**, 17 July, 267; 1879.

review article

The Loch Lomond Stadial in the British Isles

J. B. Sissons

Department of Geography, University of Edinburgh, High School Yards, Edinburgh, UK

During the last very cold period in the British Isles, which occurred more than 10,000 radiocarbon years ago, glaciers developed in many mountain areas, permafrost was extensive, frost break-up of rocks was severe (its effects including a pronounced coastal platform and cliff), downslope movement of debris was greatly accelerated, many rivers spread coarse debris over low ground, the vegetation was transformed and the fauna was greatly affected.

FOLLOWING the decay of the last ice sheet¹, oceanic polar water spread southwards again and reached the latitude of south-west Ireland². The associated cold period in the British Isles is the Loch Lomond Stadial, approximately equivalent to the Younger Dryas of Scandinavia. The effects of the stadial were probably more pronounced and varied in the British Isles than anywhere else, due partly to varied relief but particularly to maritime location and the position of the maximal extent of polar water. Norway is similarly susceptible to oceanic temperature changes but was then mostly covered by the Scandinavian ice sheet. In continental countries now bordering the southern North Sea, the maritime influence would have been diminished due to location and because this part of the North Sea was then land (because of low world sea level). Comparison of pollen-bearing Younger Dryas sediments from many sites shows that very distinctive floras and sediments occur in the British Isles, but the stadial changes are less marked in the North European Plain and diminish southwards and eastwards of it³.

Glaciers

Small cirque glaciers developed at various places in the mountains of Ireland and Wales. Their maximal limits are normally defined by fresh-looking end moraines; other end moraines frequently record subsequent intermittent retreat. In Ireland such moraines occur, for example, in the Wicklow, Mourne and Galty mountains, in the Nephin Beg range, in County Donegal and in western County Kerry⁴. In Wales they are present, for example, in the Brecknock Beacons, on Cader Idris and especially in Snowdonia⁵⁻⁸.

The 64 glaciers that developed in the Lake District included many small cirque glaciers but there were also valley glaciers up to 5 km long (Fig. 1). The limits of the majority of the glaciers are marked by end moraines, the limiting moraine being often immediately succeeded by one to three other moraines. In a few places fluted moraines record the final direction of the ice movement and in many places hummocky moraine lies within the limits of the former glaciers. Despite their considerable number, the Lake District glaciers covered an area of only 55 km² and their combined volume was only 3.3 km³. Small glaciers are said to have existed in the northern Pennines⁹ and an isolated one developed on the Cheviot. In the Southern Uplands

of Scotland, glaciers were most extensive in the centre around Broad Law but even here the largest was only a few kilometres long.

In total contrast a very large area of glacier ice accumulated in the western highlands of Scotland: its ice-shed altitude attained 850–900 m, numerous mountains (not shown in Fig. 2) rising above its surface. Over Rannoch Moor its thickness exceeded 400 m and in the southern Great Glen reached 600 m. One small part of this ice mass, the Loch Lomond Glacier, had a volume of ~80 km³, thus far exceeding the combined volume of all the glaciers mentioned above¹⁰. The numerous other glaciers that formed in the Scottish Highlands and some of the islands included many small cirque glaciers as well as valley glaciers up to 10 km long. Some glaciers developed beneath rock walls lacking cirque form and others in broad, shallow valleys in plateau surfaces. This last type of accumulation resulted in an ice cap of 65 km² in the south-east Grampians and one of 360 km² in the central Grampians¹¹⁻²⁰. Yet glaciers were not ubiquitous in the Highlands. In the Monadhliath Mountains, where a large area of ground exceeds 600 m, there were only two; in the north-west Cairngorms only three out of seven cirques at 1,000 m altitude contained glaciers¹⁹.

The landforms produced by the stadial glaciers in the Highlands are often clear and abundant. There are scores of end moraines, many of them small, but the longest extend continuously for 15 km or intermittently for 40 km, the highest has outer slopes 30–40 m high, and the broadest individual end moraine ridges are 200 m wide. These clear, often arcuate, end moraines record the maximal extent of many of the former glaciers: other clear end moraines often occur immediately behind them.

Many glaciers left extensive areas of hummocky moraine in the Scottish Highlands and some small ones deposited great quantities of boulders as continuous undulating spreads. Fluted moraines are especially common in north-west Scotland and Mull^{12,16}. Hundreds of small (1–4 m deep) meltwater channels were formed in association with the large plateau ice cap of the central Grampians¹⁴. Outwash terraces occur at and near the limits of many of the former glaciers in the central and eastern Grampians but are small compared with a buried outwash fan in the western Forth valley²¹ and with outwash masses at certain west coast localities (for example, by Lochs Etive, Linnhe and Shiel)^{22,23}. Yet the morphological evidence is not prolific everywhere: the limit of the large west Highland ice mass is

still not fully mapped partly due to the lack of clear landforms (Fig. 2).

Ice-dammed lakes

A lake held up by the Glen Moriston glacier (west of Loch Ness) occupied at least nine different levels, seven of them marked by shorelines. Excellent cross-valley moraines were formed at the ice margin during retreat. Sudden release of lake waters produced belts of rock swept free of drift that extend intermittently for several kilometres¹⁷. The 'parallel roads' of glens Roy, Gloy and Spean are mainly lake shorelines²⁴ formed by frost erosion and occur at five levels in Glen Roy²⁵. The final Roy/Spean lake had an area of 73 km², a maximal depth of 200 m and a volume of 5 km³. It has been suggested that catastrophic emptying of this lake in a few days, with peak discharge of 22,500 m³ s⁻¹, caused the surface level of Loch Ness to rise 8.5 m and resulted in a massive spread of sand and gravel at the southern end of the loch²⁶.

Periglacial activity

The very cold climate of the stadial resulted in widespread and varied periglacial activity. In the Scottish Highlands especially, powerful frost wedging produced blockfields and extensive spreads of smaller debris that now mantle many summits and plateaus. Such debris often moved down slopes as sheets with steep, often crenulate (lobate) frontal margins typically 1–5 m high, the agency probably being interstitial ice. In many places formation may have begun during ice-sheet decay but the major lobes and sheets as now widely preserved in fossil form were produced during the stadial, this being in accord with their absence from ground then occupied by glaciers.

At the base of a few slopes in the north-west Cairngorms and the Lake District accumulations of coarse debris with steep fronts up to 10 m high seem to be fossilised valley-wall rock glaciers. Fossil rock glaciers with curving terminal ridges resembling end moraines occur in the north-west Cairngorms and Jura, having been supplied with abundant coarse debris from adjacent steep slopes^{19,27}. Deposits associated with former perennial snow beds are much more common. Protalus ramparts occur in Snowdonia⁶ and one on Fan Fawr in south Wales is 11 m high⁷. They also occur in the Lake District and at many

locations in the Scottish Highlands, one in the Cairngorms having an outer slope 10–20 m high. By far the largest feature known is a complex rampart in north-west Scotland, which is 1 km long and up to 55 m high, its volume implying a 17 m retreat of the rock face that supplied it²⁸. Some snow-bed deposits comprise groups of mounds of angular debris (for example, in the Lake District) and others consist of great quantities of boulders forming spreads and ill-defined heaps; one such deposit in the Cairngorms extends for 2 km. None of the above features is firmly dated although their distribution complements that of the stadial glaciers, implying approximate contemporaneity of formation.

Fossil screes are widespread: in Wales, for example, they occur especially on south- and west-facing slopes favoured by frequent freeze-thaw cycles. The near-surface layer of angular gravel, often ~0.5 m thick, is believed to have accumulated during the Loch Lomond Stadial^{29,30}. Downslope movement of debris at low altitudes during the stadial is demonstrated in the Glasgow area, Dumfriesshire, Banffshire and the East Midlands by radiocarbon dates^{31–34}. In Edinburgh a 2 m layer of stratified stony clay sandwiched between lake deposits^{35,36} suggests stadial transfer of debris from Salisbury Crags by solifluction and snow meltwater over an average slope of 4–5°. Much further south, rapid frost break-up of the chalk occurred and trails of chalk rubble and mud extend for 2 km from coombes in the North Downs. These sheets of chalky slurry were carried by water released from melting snow and thawing of frozen ground. They often overlie an organic layer representing the Windermere Interstadial and their volume implies that at least some coombes were mainly formed during the stadial^{37,38}. Soil instability and incomplete vegetation cover are indicated widely in the British Isles by minerogenic deposits in present and former lakes (although the minerogenic layer may not correspond exactly with the stadial as indicated by pollen analysis).

Occasional thin seams of silt in chalky stadial deposits point to wind action^{37,38}. A radiocarbon date and pollen analyses show that part of the wind-blown Shirdley Hill Sand of Lancashire accumulated during the stadial³⁹. In the Vale of York and vicinity wind-blown sand occurs extensively as slightly hummocky sheets or as low dunes, radiocarbon dates demonstrating mainly stadial deposition^{40–42}. Small fossil wedges infilled with the sand suggest permafrost, as do the wedges related to the stadial scree deposits in central Wales^{29,30} and frost cracks associated with a stadial gravel layer in Buckinghamshire⁴³. In Wales organic sedimentation in features interpreted as fossil pingos began very early in the Flandrian^{44–46}, suggesting they ceased to form during the stadial, but this is not readily reconcilable with their thick infill (10 m) of clay. Pollen analysis suggests that some at least of the small (30–75 m diameter) pingos of south-east Ireland formed during the stadial⁴⁷. Some of the fossil pingos on chalk near Cambridge are believed to have existed at this time, their restriction to locations where hydrological conditions were most suitable suggesting that permafrost was discontinuous⁴⁸. Fossil frost wedges just within the Loch Lomond Advance limit at three western sites near sea level in Scotland point to widespread permafrost in that country⁴⁹.

Fluvial activity

Snowmelt and surface thawing of frozen ground together with direct precipitation caused spring floods and abundant deposition on low ground. In Edinburgh (Corstorphine Loch) an extensive alluvial fan is dated by Windermere Interstadial deposits below it and Flandrian deposits above it. The age of large fans at the mouths of small tributary valleys in Glen Roy is demonstrated by their relation to the former ice-dammed lake. The lower parts of numerous valleys cut in sandy raised beach deposits in Scotland are infilled with very early Flandrian marine deposits, implying their excavation during the stadial when subsurface drainage was prevented by permafrost⁵⁰. At the mouths of three Longmynd valleys in Shropshire fans of angular

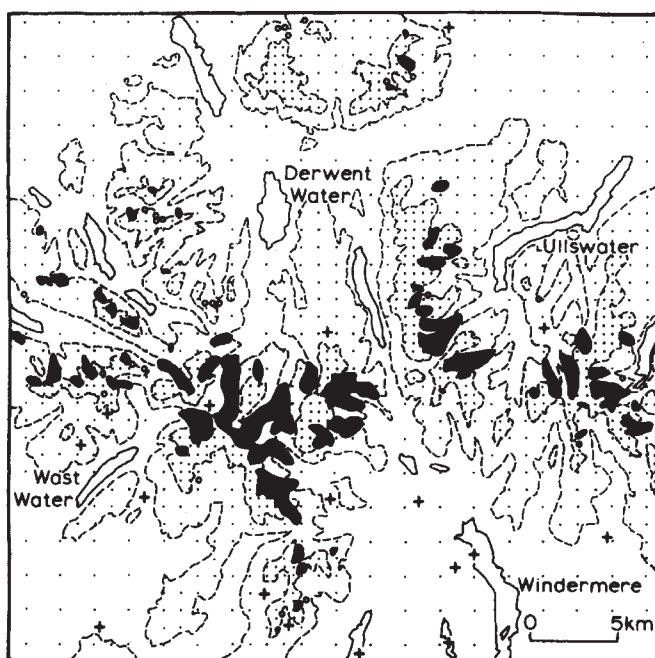
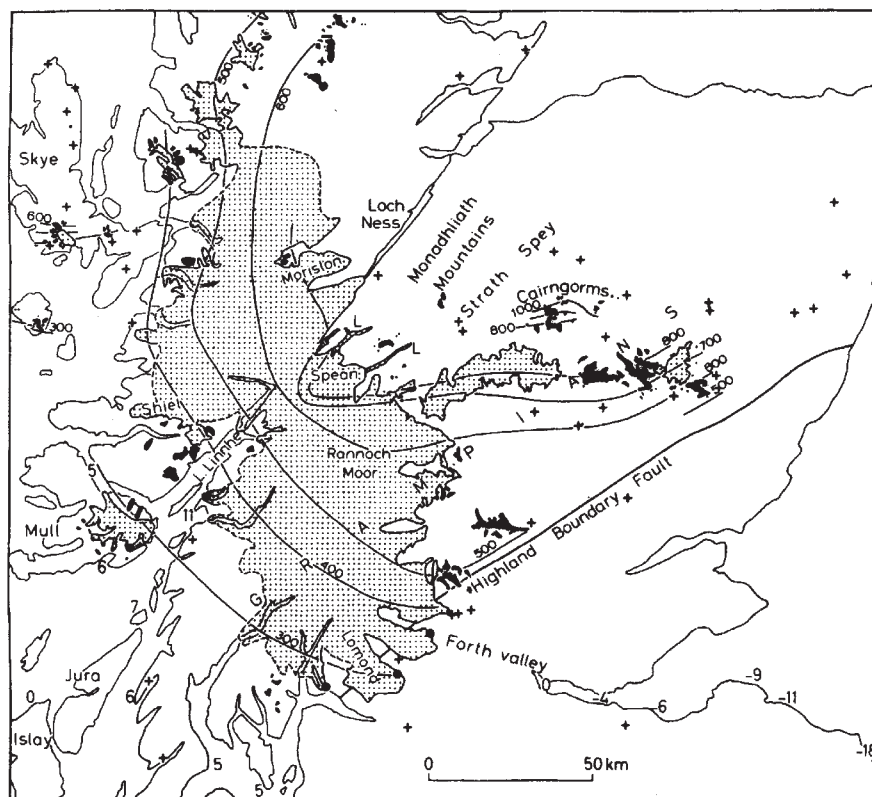


Fig. 1 Extent of stadial glaciers (solid) and locations of perennial snow beds (circles) in the Lake District. Lateglacial pollen sites shown by crosses. Altitude indicated by density of stipple, contours being at 300 and 600 m. Based on refs 103 and 110 and unpublished data.

Fig. 2 Extent of stadial glaciers in part of the Scottish Highlands (stipple and solid) and regional firn-line altitudes (100 m interval): a dashed line is used where glacier limits are not yet mapped. Radiocarbon-dated marine shell sites related to glaciers shown by dots. Late-glacial pollen sites shown by crosses. Ice-dammed lakes of the Glen Roy area shaded and marked L. Representative altitudes for the Main Lateglacial Shoreline are indicated. Based on numerous sources including unpublished data provided by H. J. B. Birks, M. Robinson, D. G. Sutherland, K. S. R. Thompson, P. Thorpe and T. Wain-Hobson.



debris, one known to overlie Windermere Interstadial deposits, were partly laid down by stadial floods⁵¹. Even in the low-amplitude relief of Suffolk radiocarbon dates and other evidence indicate that sand and gravel deposits up to 400 m wide and 12 m thick accumulated in the Gipping valley during the stadial and very early Flandrian⁵².

Flora and fauna

Open tundra vegetation prevailed during the stadial: there is only occasional macroscopic evidence suggesting the existence of tree birches, as at one site in Skye⁵³ and in the lower stadial deposits of a site in Lincolnshire⁵⁴, but dwarf birch was common. In pollen profiles annual pollen deposition rates are low and there is an increase in poorly preserved pollen. Pollen of Compositae, Cruciferae, Caryophyllaceae and Chenopodiaceae accord with open ground and disturbed soils, *Rumex* often being abundant and *Artemisia* especially characteristic. Large areas of fresh mineral soils are indicated by pollen of plants that favour base-rich conditions. The very low carbon content of stadial sediments in some lake sites in the Lake District and northern Scotland implies complete destruction of soils developed in the preceding interstadial, as well as low aquatic productivity⁵⁵⁻⁵⁷. The large numbers of moss spores in the sediments and, in northern Scotland, an increase (compared with underlying sediments) in the proportion of aerophile diatom taxa⁵⁸ are in accord with this evidence.

The crustacean *Lepidurus arcticus* has been found in stadial deposits in many parts of the British Isles. At Blelham Bog in the Lake District the arctic Cladocera *Chydorus sphaericus* and *Acroperus harpae* increased in numbers⁵⁹. Among beetles there was a preponderance of northern species, such as *Nebria nivalis* Pk, which is now found in Scotland only on the highest mountains, and *Bembidion dauricum*, now in Europe confined to arctic Norway and Swedish Lapland⁶⁰⁻⁶⁴. Vertebrates included reindeer, arctic lemming, Norway lemming, tundra vole and ptarmigan^{65,66}.

Marine activity

The stadial in some Scottish inshore waters is represented by thick deposits (such as 13 m) with foraminifera, ostracoda

and/or dinoflagellate cysts indicating cold or very cold conditions^{67,68}. These overlie interstadial 'Clyde beds', whose limited distribution above sea level is largely determined by marine erosion that occurred during the following cold period. This erosion, due to severe frost action combined with wave action, produced a platform often 20–30 m (but sometimes over 150 m) wide backed by a cliff up to 15 m high now seen on the Scottish coast between Loch Linnhe and the Firth of Clyde⁶⁹⁻⁷¹. The shoreline (Main Lateglacial Shoreline) slopes owing to glacio-isostatic recovery, having a maximal altitude of ~11 m near Oban and passing below sea level in southern Kintyre and north-east Islay. Around Grangemouth in the Forth valley 28 km² of drift and, locally, bedrock were planated⁵⁰. The shoreline is at -18 m, 9 km north of Berwick (Fig. 2), where the associated rock platform is probably 600 m wide^{72,73}. The buried channel of the Tweed may well relate to this low sea level because it descends to ~-20 m and, significantly, lacks glacial till⁷⁴. Pre- and post-stadial sea-level evidence from southern parts of England, Wales and Ireland⁷⁵⁻⁸³ suggests a stadial sea level there between -35 and -50 m. The Main Lateglacial Shoreline may be represented by the sharp break of slope traced for 30 km at -35 to -40 m in Cardigan Bay⁸⁴ and by the pronounced erosional shoreline at -42 m identified off part of the coast of south-west England^{85,86}.

Climate

All the preceding evidence is intimately linked to climate. Suggested mean July temperatures for the stadial are 10 °C for southern England and 9 °C for northern England (beetle evidence)⁶³, 7.5 °C for the Lake District⁹, 7 °C for the south-west Grampians¹⁰ and 6 °C for the south-east Grampians⁸⁷ (sea level values based on glaciers). Since discontinuous permafrost indicates a mean annual temperature no higher than -1 °C, such figures imply a greater annual temperature range than at present, as (independently) do beetle assemblages; causes include greater differences than now between amounts of winter and summer solar radiation reaching the top of the atmosphere over the British Isles⁸⁸, the smaller North Sea, probable freezing of the surface of this sea in winter, and an ice sheet over most of Scandinavia.

Glacier dimensions, altitudes and distributions imply that the principal snowfalls were associated with south to south-east air streams preceding warm or occluded fronts, although south-west air streams were important contributors of precipitation in western areas¹⁰. This suggests that many depressions followed more southerly tracks than prevail today⁸⁷, related to the junction of polar and warmer oceanic water masses at the latitude of south-west Ireland. The stormy climate resulted in extensive wind transfer of snow to glaciers and snow beds, especially by south-west winds, allowing widespread periglacial action on high ground thus largely denuded of snow. In the Scottish Highlands glacier firn-line altitudes (Fig. 2) imply that precipitation diminished inland from the west coast and northwards from the Highland Boundary Fault: in Strath Spey it may have been only 200–300 mm yr⁻¹ (ref. 10). These precipitation inferences from glaciers in the Highlands accord broadly with proportions of *Artemisia* pollen in stadal pollen counts. Species of *Artemisia* favour well-drained ground and normally do not tolerate much snow. Percentages of *Artemisia* pollen increase eastwards into the Highlands from the west coast and northwards from the Highland Boundary Fault, culminating in Strath Spey, where at one site *Artemisia* constitutes 40–65% of total land pollen almost throughout the stadial^{55,89–94}.

Dating and climatic changes

Dates of ~10,800–10,300 yr BP have often been assigned to the stadial. Many radiocarbon dates are in accord but many others are not. Dates assigned to the beginning of the stadial are normally for the uppermost organic material underlying the stadial minerogenic layer in present or former lake sites. Dates may vary because of thickness of material sampled⁹⁵, contamination or other errors (several sequences show reversals) and because the initiation of minerogenic sedimentation would have been influenced by local factors (such as slope angles, and altitude) and geographic position. Due to the lag in vegetational response, and hence in the cessation of organic sedimentation, such radiocarbon dates will tend to be younger than the causative climatic change. Similarly, dates for the end of the stadial will tend to be too young, further factors here being soil removal during the stadial and slope instability. Coarse fluvial deposits continued to accumulate in the early Flandrian in Suffolk⁵² and Shropshire⁵¹ and are attributed to the very early Flandrian in the Isle of Man⁹⁶, while wind-blown sand continued to be deposited after the end of the stadial in the Vale of York⁴¹. In the Forth valley the highly micaceous minerogenic deposits of the Main Buried Beach, which ceased to accumulate at 9,600 yr BP⁹⁷, contrast with the overlying organic carse (mud-flat) deposits. Thus different processes that characterised the stadial operated over different periods of time and responded differently to climatic change.

Nine radiocarbon dates for material picked up or over-ridden by advancing stadial glaciers all lie between 11,800 and 11,300

yr BP^{12,98–100}. Seven of these dates relate to five marine shell sites bordering the south-west Grampians (Fig. 2) and the other two are for a lake site in the Wicklow Mountains. They imply that glaciers reached their maximal extent later.

Coleoptera indicate that a decline in temperature began before 12,000 yr BP³⁴, Cladocera imply a deteriorating climate perhaps from ~11,800 BP⁵⁹, and some pollen evidence indicates irregularly declining temperatures from ~11,800 BP^{57,101}. Since only a very few degrees drop in present summer temperatures would cause glaciers to form in Scotland¹⁰², glaciers may have begun to develop here by 11,500 BP. Varied relief and precipitation and the southwards spread of polar water point to glacier initiation at different times in different places. The 400–450 varves of major Lake District lakes¹⁰³, which encompass both glacier advance and retreat (unpublished data) are far less numerous than those in Loch Ness⁵⁵, implying that glaciers existed for a much longer period in the Loch Ness catchment.

Retreat moraines show that some of the larger glaciers in the Lake District and central Southern Uplands withdrew actively (unpublished data). However, most Lake District glaciers and some Southern Uplands ones produced end moraines only at and near their furthest limits (in a zone 50–150 m wide). In the Scottish Highlands, except at ice-dammed lake sites and possibly at one other site, clear end moraines are restricted to the same situation, the end moraine zone being several hundred metres wide for some of the larger glaciers. It is tempting to suggest that all these end moraine zones are synchronous and that subsequently most glaciers decayed rapidly owing to the rapid rise in summer temperatures demonstrated by Coleoptera¹⁰⁴, this in turn being related to withdrawal of polar water (or to the latter having been rapidly over-ridden by warmer water). However, there is insufficient evidence to test this hypothesis.

Basal organic sediments from a site located behind a major end moraine at Callander have been dated as 10,670±85 yr BP⁹³. Similar sediments from three sites on Rannoch Moor, in the heart area of the major stadial ice mass, range between 10,390±200 yr BP and 10,660±240 yr BP^{105–108}. If valid these dates imply much earlier deglaciation than is normally assumed. It has been suggested that the dates may be too old due to the unique chemical and hence biological character of recently deglaciated terrain¹⁰⁹. Yet this explanation does not apply to the Lake District evidence: here glaciers had withdrawn from the end-moraine belt before the end of the *Artemisia* pollen assemblage zone (ref. 103 and unpublished data) which, according to a radiocarbon date from Blenheim Bog (located outside the glacier limits), ended before 10,650±170 yr BP¹⁰¹. This evidence suggests the attainment of glacier maxima ~10,800 yr BP, markedly out of phase with much other evidence. Many more radiocarbon dates are needed to resolve this apparent discrepancy.

I thank many present and former research students for comments.

- Sissons, J. B. & Walker, M. J. C. *Nature* **249**, 822–824 (1974).
- Ruddiman, W. F., Sancetta, C. D. & McIntyre, A. *Phil. Trans. R. Soc. B* **280**, 119–142 (1977).
- Watts, W. A. *Quat. Newslett.* **27**, 5 (1979).
- Davies, G. L. H. & Stephens, N. *The Geomorphology of the British Isles: Ireland* (Methuen, London, 1978).
- Seddon, B. J. *Glaciol.* **3**, 94–99 (1957).
- Unwin, J. D. *Cambria* **2**, 20–33 (1975).
- Lewis, C. A. *Brycheiniog* **14**, 97–120 (1970).
- Ellis-Griffith, I. D. *Cambria* **4**, 46–55 (1977).
- Manley, G. *Liverp. Manch. geol. J.* **2**, 188–215 (1959).
- Sissons, J. B. *Nature* **278**, 518–521 (1979).
- Ballantyne, C. K. & Wain-Hobson, T. *Scott. J. Geol.* (in the press).
- Gray, J. M. & Brooks, C. L. *Scott. J. Geol.* **8**, 95–103 (1972).
- Sissons, J. B. *Scott. geogr. Mag.* **88**, 168–181 (1972).
- Sissons, J. B. *Trans. Inst. Br. Geogr.* **62**, 95–114 (1974).
- Sissons, J. B. *Scott. J. Geol.* **13**, 23–36 (1977).
- Sissons, J. B. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 45–59 (Pergamon, Oxford, 1977).
- Sissons, J. B. *Trans. Inst. Br. Geogr.* **2**, 224–242 (1977).
- Sissons, J. B. *Scott. J. Geol.* (in the press).
- Sissons, J. B. *Scott. geogr. Mag.* (in the press).
- Sissons, J. B. & Grant, A. J. H. *Scott. J. Geol.* **8**, 85–93 (1972).
- Smith, D. E., Thompson, K. S. R. & Kemp, D. D. *Boreas* **7**, 97–107 (1978).
- McCann, S. B. *Scott. J. Geol.* **2**, 84–95 (1966).
- Gray, J. M. *Proc. Geol. Ass.* **86**, 227–238 (1975).
- Jamieson, T. F. *Quart. J. geol. Soc.* **19**, 235–259 (1863).
- Sissons, J. B. *Boreas* **7**, 229–244 (1978).
- Sissons, J. B. *J. geol. Soc. Lond.* (in the press).
- Dawson, A. G. *Scott. J. Geol.* **13**, 37–41 (1977).
- Sissons, J. B. *Scott. geogr. Mag.* **92**, 182–190 (1976).
- Watson, E. *Biul. Peryglac.* **18**, 95–113 (1969).
- Potts, A. S. *Geogr. Ann.* **53**, 39–51 (1971).
- Godwin, H. & Willis, E. H. *Radiocarbon* **1**, 63–75 (1959).
- Harkness, D. D. & Wilson, H. W. *Radiocarbon* **16**, 238–251 (1974).
- Dickson, J. H., Jardine, W. G. & Price, R. J. *Nature* **262**, 43–44 (1976).
- Bishop, W. W. & Coope, G. R. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 61–88 (Pergamon, Oxford, 1977).
- Taylor, A. *Trans. Edinb. geol. Soc.* **5**, 44–48 (1888).
- Cadell, H. M. *Trans. Edinb. geol. Soc.* **6**, 287–296 (1893).
- Kerney, M. P. *Phil. Trans. R. Soc. B* **246**, 203–254 (1963).
- Kerney, M. P., Brown, E. H. & Chandler, T. J. *Phil. Trans. R. Soc. B* **248**, 135–204 (1964).
- Tooley, M. J. *The Isle of Man, Lancashire Coast and Lake District*, X INQUA Cong. Excursion Guide (1977).
- Matthews, B. *Nature* **227**, 1234–1236 (1970).
- Gaunt, G. D., Jarvis, R. A. & Matthews, B. *Proc. Yorks. geol. Soc.* **38**, 281–284 (1971).
- Shotton, F. W. & Williams, R. E. G. *Radiocarbon* **15**, 451–468 (1973).
- Evans, J. G. *Proc. Geol. Ass.* **77**, 347–364 (1966).
- Watson, E. *Geol. J.* **7**, 381–392 (1971).
- Handa, S. & Moore, P. D. *New Phytol.* **77**, 205–225 (1976).
- Watson, E. *Phil. Trans. R. Soc. B* **280**, 183–198 (1977).
- Mitchell, G. F. *Proc. R. Irish Acad. B* **73**, 269–282 (1973).
- Sparks, B. W., Williams, R. B. G. & Bell, F. G. *Proc. R. Soc. Lond. A* **327**, 329–343 (1972).
- Sissons, J. B. *Scott. J. Geol.* **10**, 311–337 (1974).
- Sissons, J. B. *Trans. Inst. Br. Geogr.* **48**, 19–50 (1969).
- Rowlands, P. H. & Shotton, F. W. *J. geol. Soc. Lond.* **127**, 599–622 (1971).

52. Rose, J., Turner, C., Coope, G. R. & Bryan, M. D. in *Timescales in Geomorphology* (eds Cullingford, R. A., Davidson, D. A. & Lewin, J.) (Wiley, New York, in the press).
53. Birks, H. J. B. *The Past and Present Vegetation of the Isle of Skye: a Palaeoecological Study* (Cambridge University Press, 1973).
54. Suggate, R. P. & West, R. G. *Proc. R. Soc. B* **150**, 263–283 (1959).
55. Pennington, W., Haworth, E. Y., Bonny, A. P. & Lishman, J. P. *Phil. Trans. R. Soc. B* **264**, 191–294 (1972).
56. Pennington, W. in *Palaeolimnology of Lake Biwa and the Japanese Pleistocene* (ed. Horie, S.) 536–569 (1975).
57. Pennington, W. *Phil. Trans. R. Soc. B* **280**, 247–271 (1977).
58. Haworth, E. Y. *New Phytol.* **77**, 227–256 (1976).
59. Harmsworth, R. V. *Ecol. Mon.* **38**, 223–241 (1968).
60. Coope, G. R., Morgan, A. & Osborne, P. J. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **10**, 87–101 (1971).
61. Osborne, P. J. *Proc. Croydon nat. Hist. Sci. Soc.* **14**, 162–175 (1971).
62. Coope, G. R. in *Ice Ages: Ancient and Modern* (eds Wright, A. E. & Moseley, F.) 153–168 (Liverpool, 1975).
63. Coope, G. R. *Phil. Trans. R. Soc. B* **280**, 313–340 (1977).
64. Osborne, P. J. *Phil. Trans. R. Soc. B* **263**, 327–367 (1972).
65. Allison, J., Godwin, H. & Warren, S. H. *Phil. Trans. R. Soc. B* **236**, 169–240 (1952).
66. Stuart, A. J. *Phil. Trans. R. Soc. B* **280**, 295–312 (1977).
67. Binns, P. E., Harland, R. & Hughes, M. J. *Nature* **248**, 751–754 (1974).
68. Peacock, J. D. *Bull. geol. Surv. Gt. Br.* **48**, 55–67 (1974).
69. Gray, J. M. *Trans. Inst. Br. Geogr.* **61**, 81–99 (1974).
70. Sissons, J. B. *Boreas* **3**, 41–48 (1974).
71. Gray, J. M. *Trans. Inst. Br. Geogr.* **3**, 151–164 (1978).
72. Eden, R. A., Carter, A. V. F. & McKeown, M. C. *Mar. Geol.* **7**, 235–251 (1969).
73. Sissons, J. B. *Scott. geogr. Mag.* **92**, 17–29 (1976).
74. Rhind, D. W. *Trans. nat. Hist. Soc. Northumb.* **4**, 159–164 (1972).
75. Godwin, H. & Willis, E. H. *Radiocarbon* **6**, 116–137 (1964).
76. Kidson, C. & Heyworth, A. *Proc. Usher Soc.* **2**, 565–584 (1973).
77. Stillman, C. J. *Sci. Proc. R. Dubl. Soc. A* **3**, 125–127 (1968).
78. Clarke, R. H. *Quart. J. geol. Soc.* **125**, 277–318 (1970).
79. Shotton, F. W., Williams, R. E. G. & Johnson, A. S. *Radiocarbon* **16**, 285–303 (1974).
80. Shephard-Thorn, E. R. & Wymer, J. J. *South-East England and the Thames Valley*, X INQUA Cong. Excursion Guide (1977).
81. Kidson, C. & Heyworth, A. *Nature* **273**, 748–750 (1978).
82. Brown, A. P. *New Phytol.* **78**, 525–531 (1977).
83. Evans, D. J. & Thompson, M. S. *Proc. Geol. Ass.* **90**, 1–14 (1979).
84. Dobson, M. R., Evans, W. E. & James, K. H. *Mar. Geol.* **11**, 27–69 (1971).
85. Cooper, L. H. N. *Nature* **161**, 280 (1948).
86. Kelland, N. C. *J. geol. Soc. Lond.* **131**, 7–17 (1975).
87. Sissons, J. B. & Sutherland, D. G. *J. Glaciol.* **17**, 325–346 (1976).
88. Vernekar, A. D. *Meteorol. Mon.* **12** (1971).
89. Walker, M. J. C. *Pollen Spores* **17**, 67–92 (1975).
90. Walker, M. J. C. *J. Biogeogr.* **2**, 265–284 (1975).
91. Pennington, W. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 119–141 (Pergamon, Oxford, 1977).
92. Lowe, J. J. & Walker, M. J. C. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 101–118 (Pergamon, Oxford, 1977).
93. Lowe, J. J. *Pollen Spores* **20**, 367–397 (1978).
94. Birks, H. H. & Mathewes, R. W. *New Phytol.* **80**, 455–484 (1978).
95. Gray, J. M. & Lowe, J. J. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 163–181 (Pergamon, Oxford, 1977).
96. Mitchell, G. F. Q. *Jl geol. Soc.* **121**, 359–381 (1965).
97. Sissons, J. B. *Trans. Inst. Br. Geogr.* **39**, 19–29 (1966).
98. Sissons, J. B. *Scott. J. Geol.* **3**, 375–381 (1967).
99. Peacock, J. D. *Nature phys. Sci.* **230**, 43–45 (1971).
100. Huddart, D. *South-East Ireland*, X INQUA Cong. Excursion Guide (1977).
101. Pennington, W. *Boreas* **4**, 157–171 (1975).
102. Manley, G. *Geogr. Ann.* **31**, 179–193 (1949).
103. Pennington, W. in *Geology of the Lake District* (ed. Moseley, F.) 207–225 (1978).
104. Coope, G. R. & Brophy, J. A. *Boreas* **1**, 97–142 (1972).
105. Lowe, J. J. & Walker, M. J. C. *Nature* **264**, 632–633 (1976).
106. Walker, M. J. C. & Lowe, J. J. *J. Biogeogr.* **4**, 333–351 (1977).
107. Walker, M. J. C. & Lowe, J. J. *J. Biogeogr.* (in the press).
108. Walker, M. J. C. & Lowe, J. J. in *Timescales in Geomorphology* (eds Cullingford, R. A., Davidson, D. A. & Lewin, J.) (Wiley, New York, in the press).
109. Sutherland, D. G. *Quat. Newslett.* **27**, 11–12 (1979).
110. Walker, D. *New Phytol.* **54**, 222–254 (1955).

articles

General relativity, thermodynamics, and the Poincaré cycle

Frank J. Tipler

Department of Mathematics, University of California at Berkeley, Berkeley, California 94720

An arbitrarily close return to a previous initial state of the Universe, such as predicted by the Poincaré recurrence theorem, cannot occur in a closed universe governed by general relativity. The significance of this result for cosmology and thermodynamics is discussed.

THE idea that history repeats itself—that every state of the Universe has occurred before and will occur again *ad infinitum*—is a concept which originated thousands of years ago and in the form of the Poincaré cycle and the ‘cyclic’ universe, still plays an important part in fundamental physics. This notion of an ‘eternal return’ will be considered here in the context of general relativity, and it will be shown that in a closed universe, no arbitrarily close (or exact) return to a previous state of the universe is possible, provided certain very general restrictions on the matter tensor and global causal structure hold. This result is in stark contrast to the situation in classical mechanics: Poincaré showed^{1,2} that for almost all initial states, any mechanical system with a finite number of degrees of freedom, finite total and kinetic energy, and constrained to move within a finite box must return arbitrarily closely and infinitely often to almost every previous state of the system. The reason for ‘no return’ in general relativity and ‘eternal return’ in classical mechanics is that in general relativity, singularities intervene to prevent recurrence. General-relativistic closed universes are thought to begin and end in singularities of infinite curvature,

and these singularities force time in general relativity to be linear rather than cyclic.

Poincaré recurrence has been reported^{2–5} to be a major stumbling block in the definition of entropy, a function of the state of the system which never decreases, in terms of the fundamental microscopic variables of the system. Hitherto the increase of an ‘entropy’ defined by such variables has been obtained (not very successfully) by such tricks as ‘coarse-graining’^{2,6}; an *ad hoc* addition of randomness to the system in the form of a postulate of ‘molecular chaos’ or ‘random phases’^{2,3}; or by taking the thermodynamic limit^{7,8}. The thermodynamic limit, which involves letting the volume of the system and the number of particles it contains go to infinity, is particularly unphysical as the infinite volume limit places one in the realm of cosmology, yet the limit ignores the most important cosmological force, gravity. The results of this paper may make such processes unnecessary, and there may be a strong connection between the increase of gravitational field entropy, which does not seem to require the above tricks, and the increase of matter entropy. The paper concludes with a discussion of the significance of the no-return theorem for cosmology.

To prove that no two states of the Universe can be identical or even arbitrarily close, we must first define the notion of ‘close’. This can be done by regarding the set of all initial data as a Sobolev space W^s . The global topology of the space-time (M, g) is $S \times R$, where S is compact, if the space-time is globally hyperbolic. We choose some positive definite metric e_{ab} on M ,

and define a norm on W^* by

$$\|K\|_m = \left[\sum_{p=0}^m \int_S (|D^p K|)^2 d\sigma \right]^{1/2}$$

$$\|h, \chi\| = \|h\| + \|\chi\| \quad (1)$$

where $d\sigma$ is the volume element induced on S by e_{ab} , D^p is the generalised p th covariant derivative with respect to some chosen background metric $\bar{g}_{ab}$, and single verticle lines denote is the norm induced by e_{ab} (see ref. 9, p 233–235 for more details). Two tensors will be 'close' if they are 'close' in the norm of equation (1).

Assumptions in the no-return theorem

The no-return theorem will require four conditions on the matter tensor. The first condition, the time-like convergence condition (ref. 9, p 95), says roughly that gravitation is always an attractive force. The other three conditions, are here called (a), (b) and (c) (for precise definition see ref. 9, p 254–255). Briefly, (a) says that the development from a given set of initial data is locally unique; (b) says that this unique development is locally stable; and (c) says that the stress-energy tensor is a polynomial in the matter fields, their first derivatives, and the space-time metric. As discussed by Hawking and Ellis (ref. 9, p 255), these are physically reasonable conditions to impose on the matter fields, though strictly speaking they are not necessary conditions. Any other conditions which imply global Cauchy stability and uniqueness (as defined in ref. 9 Ch. 7) would be sufficient to prove the theorem below.

We also need two global conditions. The first, the requirement of unique development from Cauchy surfaces (ref. 9, p 205), says that laplacian determinism holds for the space-time. The second, the generic condition (ref. 9, p 101) says that every causal geodesic experiences a tidal force at least once in its history. As this would be expected to be false only for a 'measure-zero' set of solutions to the Einstein equations, the generic condition in the theorem below is analogous to the 'measure-zero' qualifications in the Poincare recurrence theorem. A closed universe is defined to be a space-time in which the Cauchy surfaces are compact. A space-time that contains two disjoint space-like Cauchy surfaces which are isometric in their initial data is said to be time periodic. (Because if the evolution equations are well-posed, the development of the second Cauchy surface will be identical to that of the first, leading to a third identical surface, and so on.)

No-return theorem

Theorem: If a space-time (M, g) containing compact Cauchy surfaces is uniquely developed from initial data on any of its Cauchy surfaces, and if (M, g) also satisfies both the generic and the time-like convergence conditions, then the space-time cannot be time periodic. Furthermore, if the matter fields ψ and their first derivatives ψ' also satisfy conditions (a), (b) and (c), then for any neighbourhood U of any Cauchy surface S_1 , a number $\varepsilon > 0$ exists such that $\|(h, \chi, \psi, \psi') - (h_1, \chi_1, \psi_1, \psi'_1)\|_{a+5} > \varepsilon$ for the initial data on any Cauchy surface S with $U \cap S$ empty. (ε Depends on e_{ab} , $\bar{g}_{ab}$ and U . The Cauchy surfaces S and S_1 are assumed space-like.)

The proof of this theorem will be published elsewhere. Note that the compactness of the Cauchy surfaces—the requirement of a closed universe—is an essential condition. If the word 'compact' is removed, then the theorem is false; it does not apply to open universes. For example, a static stop in asymptotically flat space satisfies all the conditions of the theorem except compactness, and this spacetime can be foliated by Cauchy hypersurfaces normal to the timelike killing vector field. The initial data is the same on each of these hypersurfaces.

Significance of no-return theorem

One might think that any field theory in euclidean space would have a non-recurrent property similar to the one demonstrated above, as a continuous field has an infinite number of degrees of

freedom. From a physical point of view, this is not the case. It is not possible to make a precise measurement of the field variables at every point, and, in practice, a field restricted to a finite region S would be approximated by dividing up S into a finite number of subregions, and the field in each subregion replaced by its average value in that subregion. Evolution would be via the differential field equations, but one would compare the 'average' values.

Comparing initial data sets via the Sobolev norm is analogous to comparing the average values of the field variables (and their derivatives) at one time with the average values at another time. With the Sobolev norm, one essentially takes the absolute square average value of the initial data over the entire Cauchy surface rather than dividing up the Cauchy surface into subregions; the Sobolev average is a coarser average. The arbitrary metrics e_{ab} and $\bar{g}_{ab}$ in the Sobolev norm are analogous to an arbitrary coordinate system with respect to which averages are calculated. In the classical mechanics of fields in a finite box, one can define a similar norm $\| \psi \| = [\int |\psi|^2 d\sigma]^{1/2}$ on the classical field ψ . The sum $\sum_{i=1}^n \| \psi^{(i)} \|$ with $\psi^{(i)}$ being the i th derivative of ψ , is essentially the same as the Sobolev norm. This sum is a map f from the space of all values of the field and its first n derivatives into R . If the range of f is bounded—as it must be if the solutions ψ are stable—then the evolution of ψ must result in accumulation points in the range of f . An analogous accumulation is impossible in general relativity, as the above theorem shows. In general relativity the range of f is not bounded—singularities develop in space-time (this follows from the hypotheses of the above theorem and theorem 2 of ref. 9, p 266). This sort of singular behaviour is typical of many non-linear field theories^{10,11}, but in linear field theories such as electromagnetism the solutions are stable and are hence bounded. For example, the solutions to the Schrödinger equation in a finite box must recur^{12,13} in the above average sense: for any initial value $\psi(t_0)$ for the probability distribution and any ε , there exists a time t for which $\| \psi(t) - \psi(t_0) \| < \varepsilon$.

In a finite universe without singularities there should be a finite number of physically distinguishable states (this should also be true if the matter in the universe is in the form of fields, for reasons stated above). I would expect that the states of such a universe would evolve in general in a quasi-ergodic way^{14,15}. If the evolution of the physically distinguishable states is indeed quasi-ergodic, then in infinite time, every state would with high probability occur an infinite number of times.

Another way to model the evolution of such a finite-state always-existing universe, is by regarding the evolution as a finite discrete Markov chain¹⁶ with stationary transition probabilities. That is, we will sample the state of the Universe at definite time intervals Δt , and we will assume that the state of the Universe at time t_i is determined by the previous state at time t_{i-1} and the probability matrix of going from this previous state to any other state. If the Universe has existed for an infinite time, then with probability one the states of the Universe form a closed set (ref. 16, p 384); that is, we may consider without loss of generality that the Markov chain representing the Universe is now irreducible. By theorem 4 of ref. 16 (p 392), it follows that with probability one, all states recur in the future.

Similar arguments will imply some form of recurrence in closed universes which avoid singularities by 'bouncing' at some small radius (such behaviour could, of course, occur only if some of the conditions in the no-return theorem are violated¹⁷). One could envisage the physical constants of the Universe, such as the elementary particle masses, the coupling constants, the specific entropy per baryon, and so forth, changing in some fashion at each bounce. These physical constants can be regarded as additional variables in the initial data. If the number of such constants is finite and if their range of variation is finite, then in a finite universe one would expect a finite number of physically distinguishable states. This implies accumulation points in the state space with an infinite number of bounces; with quasi-ergodic or markovian evolution, we would have recurrence of all states with high probability. Quantum-mechanical

considerations do not substantially alter this conclusion, as long as the number of physically distinguishable states is to be finite.

It is, of course, possible to avoid the recurrence conclusion by assuming that the range of variation of the physical constants is not bounded, or that the Universe, although closed, increases its radius at maximum expansion with each bounce. (Tolman¹⁸, for example, argued that the monotone increase of entropy required a monotone increase of maximum radius at each bounce.) However, if the range of physical constant variation is not bounded, then a sequence of cycles would exist such that at least one physical constant would diverge in the limit. I would regard such a divergence as a singularity 'at infinity'. If the maximum radius increases monotonically with time, then with an infinite number of bounces in the past, the sequences of values of the maximum radius must have a limit point in the past. If its limiting value is non-zero, then one must have recurrence in the past, though this 'recurrence' may take the form of a gradual cessation of change, as in the Eddington-Lemaître universe¹⁹. If its limiting value is zero, then the singularity has not been really removed, it has just been placed at temporal infinity.

But the singularities in the single-cycle closed universe can also be regarded as at temporal infinity. In fact, both Milne²⁰ and Misner²¹ have contended that because physical changes occur with diverging rapidity (as measured in proper time) near a singularity, the singularity should be regarded as occurring at temporal infinity as measured in physical time. Barrow and myself²² have pointed out that extrinsic time²³, which is a naturally defined absolute time in closed universes²⁴, has just this property of placing the singularities at temporal infinity. In the extrinsic time scale, a closed universe has existed and will exist forever; the question of what preceded the big bang does not arise.

Because of Poincaré recurrence, it is impossible to define a monotonically-increasing entropy (or any monotonically-increasing function) for non-gravitational fields or systems of particles in terms of the phase space variables of the fields or particles. However, since Poincaré recurrence does not hold for gravitational fields, it is possible to define functions which increase monotonically with time—such functions are called Lyapunov functions^{7,8}—on the phase space of the gravitational field. In fact, the extrinsic time, which is proportional to the trace of the extrinsic curvature of a leaf of a constant mean curvature Cauchy hypersurface foliation of space-time, is just such a function, as the extrinsic curvature is essentially the momentum of the gravitational field²⁵ and the extrinsic time increases monotonically in closed universes, ranging from $-\infty$ at the initial singularity to $+\infty$ at the final singularity²⁴. In non-singular

asymptotically flat space, a foliation of constant mean curvature Cauchy surfaces would have to have extrinsic time constant ($=0$). This supports the conclusion of classical mechanics that isolated systems cannot define a time direction.

Bohr²⁶ together with Prigogine and his coworkers^{4,8} have suggested on the basis of Poincaré recurrence that thermodynamic concepts such as temperature and entropy are complementary to the phase space variables. In other words, they argue that a detailed microscopic description of the behaviour of a physical system would preclude the definition of thermodynamic variables for that system. If the gravitational field (or more precisely, the global structure of space-time) is taken into account, we see that it becomes possible in principle to define the thermodynamic variables and the phase space variables in a closed universe simultaneously; the notion of complementarity becomes unnecessary. This leads to speculation about the connection between quantum mechanics and the global structure of space-time: perhaps the impossibility of giving a detailed deterministic microscopic description of a quantum mechanical system is due to neglect of the global structure of space-time.

Received 19 January; accepted 3 May 1979.

1. Kac, M. *Probability and Related Topics in Physical Sciences* (Interscience, London, 1959).
2. Davies, P. C. W. *The Physics of Time Asymmetry* (University of California Press, Berkeley, 1975).
3. Brush, S. G. *The Kind of Motion We Call Heat: A History of the Kinetic Theory of Gases in the 19th Century*, Vol. 2 (North-Holland, Amsterdam, 1976).
4. Prigogine, I., George, C., Henin, F. & Rosenfeld, L. *Chem. Scripta* **4**, 5 (1973).
5. Prigogine, I., Mayné, E., George, C. & de Hahn, M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4152 (1977).
6. Uhlenbeck, G. E. in *The Physicist's Conception of Nature* (ed. Mehra, J.) 501 (Reidel, Dordrecht, 1973).
7. Prigogine, I. *Science* **201**, 777 (1978).
8. Misra, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1627 (1978).
9. Hawking, S. W. & Ellis, G. F. R. *The Large Scale Structure of Spacetime* (Cambridge University Press, 1973).
10. Glassey, R. T. *J. math. Phys.* **18**, 1794 (1977).
11. Levine, H. *Trans. Am. math. Soc.* **192**, 1 (1974).
12. Percival, I. C. *J. math. Phys.* **2**, 235 (1961).
13. Bocchieri, P. & Loinger, A. *Phys. Rev.* **107**, 337 (1957).
14. Farquhar, I. E. *Ergodic Theory in Statistical Mechanics* (Interscience, London, 1964).
15. Lebowitz, J. L. & Penrose, O. *Phys. Today* **26**, 23 (1973).
16. Feller, W. *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd edn (Wiley, New York, 1968).
17. Tipler, F. J. *Phys. Rev. D* **17**, 2521 (1978).
18. Tolman, R. C. *Relativity, Thermodynamics, and Cosmology* (Oxford University Press, 1934).
19. Tipler, F. J., Clarke, C. J. S. & Ellis, G. F. R. *Singularities and Horizons*, in *Einstein Centenary Volume* (ed. Held, A.) (in the press).
20. Milne, E. A. *Modern Cosmology and the Christian Idea of God* (Oxford University Press, 1952).
21. Misner, C. W. *Phys. Rev.* **186**, 1328 (1969).
22. Barrow, J. A. & Tipler, F. J. *Nature* **276**, 453 (1978).
23. York, J. W. *Phys. Rev. Lett.* **28**, 1082 (1972).
24. Marsden, J. E. & Tipler, F. J. preprint (1978).
25. Misner, C. W., Thorne, K. S. & Wheeler, J. A. *Gravitation* (Freeman, San Francisco, 1973).
26. Bohr, N. *J. chem. Soc.* 349 (1932).

Geochemical problems of the Antarctic dry areas

A. T. Wilson*

Chemistry Department and Antarctic Research Unit, University of Waikato, Hamilton, New Zealand

The dry areas of the Antarctic are extreme deserts. Because of the cold they are not subjected to rainfall and geochemical separations can take place which would be destroyed by flowing water in the more temperate deserts. The distribution of soluble salts which accumulate on the surface and in the lakes is unusual and is explained on the basis of relative humidity separation mechanism. It is suggested that similar phenomena may take place on Mars.

* Present address: Duval Corporation, 4715 Fort Lowell Road, Tucson, Arizona 85712.

ALTHOUGH much of the Antarctic continent is covered with ice and snow, there are some areas which are ice free. These so-called 'dry areas' are extreme deserts, characterised by extreme cold and aridity. They are also unique among deserts in being completely rainless. All precipitation falls as snow which is largely lost by sublimation. The salt geochemistry of water soluble salts which accumulate on the surface and concentrate in the lakes and ponds of these dry areas is described here. A relative humidity separation mechanism is proposed that can only operate in rainless deserts.

Although the salt and lake geochemistry of the McMurdo dry area has been studied for almost 20 years, the geochemistry of this region is far from being understood and there is a lot of

apparently conflicting evidence. Even the origin of the salts in the lakes is in dispute. Since chloride is a minor constituent of Antarctic rocks and very little chemical weathering takes place, one might expect the chloride in the lakes (which are mostly chlorides of Ca, Mg and Na) to have originated from the sea. However, some lakes such as Lake Vanda, are essentially calcium chloride brines, a fact that is difficult to explain by normal geochemical models. Further, Jones and Faure¹ have presented isotopic evidence to suggest that most of the strontium in Lake Vanda does not have an oceanic origin. Other problems are that adjacent lakes frequently have very different chemical compositions which often do not seem to match that of their inflow streams. Furthermore, different layers of the same lake sometimes have different ratios of their ionic constituents. The bottoms of some chemically stratified lakes, which on other evidence must have been out of contact with the atmosphere for hundreds if not thousands of years, contain measurable amounts of tritium (half life = 12.3 yr) in their bottom waters. Glaciers which would be expected to have the same composition as the regional snowfall, because polar glaciers do very little cutting, frequently produce meltwater streams with very different chemical compositions to the snow which falls on their neves.

This article describes phenomena which can occur only in polar deserts and presents evidence in support of an hypothesis which could explain all the apparently anomalous data. First, it is necessary to describe the Antarctic dry areas in detail.

Geochemistry of the Antarctic dry areas

Why are the dry areas not ice-covered like the rest of the Antarctic continent? In this cold and arid region (mean annual temperatures of the McMurdo oasis region are approximately -20°C) only a very small fraction of the snow that falls ever melts and almost all of it is lost directly by sublimation. The dry areas are snow free because ablation processes in total exceed all the accumulation processes. There must also be favourable topography to prevent ice from regions where accumulation processes exceed ablation from flowing into and inundating the dry area. The imaginary line which divides the region of net positive accumulation from the region where ablation exceeds accumulation is called the 'snow line'². For a given region the net precipitation increases as altitude is increased. As one moves from the sea inland the snow line rises; presumably because total precipitation decreases. As one descends below the snow line the environment becomes more and more arid. Parts of the McMurdo dry valleys are so dry that $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$ crystallises from ice-free saline ponds (such as the Don Juan Pond^{3,4}). The fact that such a deliquescent salt can exist demonstrates that this area is among the most arid on the Earth.

If one digs into the loose 'soil' of a dry area in this cold and arid region, one passes through loose dry debris and can suddenly strike a very hard and impenetrable layer, which on examination proves to be debris as above but cemented with ice crystals. This phenomenon is equivalent to the water table of the more temperate regions of the world and is called here 'the frozen water table'. The depth to this frozen water table seems to be related to distance below snow line and is probably a balance between distillation of water downwards during brief periods of snow cover and evaporation outwards during periods of low relative humidity at the surface. Certainly as one goes to higher altitudes and hence nearer to the snow line, the depth to the frozen water table decreases. During the summer, cold saline water can be found moving downslope along the surface of the frozen water table.

Drainage basins

The dry areas of Antarctica frequently consist of a number of enclosed drainage basins. The lowest part of many of these is occupied by a lake, many of which are saline but some are relatively fresh. One of the most puzzling geochemical problems of the Antarctic dry areas is that these lakes have very different chemical compositions. Data on some are given in Table 1. The

Table 1 Chemical composition of water samples from the McMurdo dry valley area of Antarctica

	$\text{Ca}^{2+}/\text{Na}^{+}$
Inland snow*	0.13
Groundwater†	1.8
Canopus Pond	0.1
Lake Fryxell	0.4
Lake Hoare	0.09
Lake Joyce	0.03
Lake Miers (bottom)	5.0
Lake Vanda	
—upper water (13 m)	1.5
—lower water (61 m)	5.4
Don Juan Pond	10.2
'Amber Ice' at base of Meserve Glacier ¹³	1.4
Seawater (for comparison)	0.04

* For example, Ice from Upper Taylor Glacier.

† For example, from surface of frozen water table east of Don Juan Pond.

lakes seem to be of three main types; (1) Lakes which have derived their salts directly from seawater flooding. These are characterised by high $\text{Mg}^{2+}/\text{Ca}^{2+}$ ratios. (2) 'Sodium chloride lakes' where the principal anion is chloride and the principal cation is sodium. These have a composition which suggests that they have been derived from the concentration of large quantities of glacial meltwater. This type of lake has a $\text{Ca}^{2+}/\text{Na}^{+}$ ratio ~ 0.1 . (3) 'Calcium chloride lakes' where the principal cations are calcium and magnesium and the principal anion chloride, the $\text{Ca}^{2+}/\text{Na}^{+}$ ratio being >1 . The calcium chloride lakes are generally relatively low in sulphate. The chemical composition of these lakes is perhaps the most difficult to understand.

Examples of lakes with salts of different origin are found in the lakes and ponds of the central Wright Valley. Canopus Pond and the upper waters of Lake Vanda have compositions of the sodium chloride type. However, the adjacent Don Juan Pond and the bottom waters of Lake Vanda have a very different composition, having a much greater proportion of calcium and magnesium chloride (see Table 1).

If the surface area of the glacier is insufficient to balance the ablation/precipitation budget, the glacier will advance further and further below the snow line until the total positive net accumulation above the snow line is balanced by the total ablation below the snow line. Usually the glacier has pushed sufficiently far below the snow line so that some summer melting takes place. In such cases, for a few days during the hottest part of the summer, a stream flows away from the glacier snout and feeds a lake which occupies the lowest point of that particular enclosed drainage basin or flows into the sea. In enclosed drainage basins the size of the lake is determined by that area needed to balance the evaporation/precipitation equation for that particular drainage area. If there is a net precipitation increase to the area the lake levels will rise and if there is a decrease in precipitation the lake levels will fall, provided that the temperature regime always remains similar.

Many of the lakes that occupy the lowest part of the various enclosed drainage basins are chemically stratified and these chemical concentration gradients may contain palaeoclimatic information.

Turning to the problem of the origin of chemical gradients in the saline lakes of the McMurdo oasis, an extreme example is Lake Vanda. The lake is 66 m deep, the lower 20 m being strongly density stratified with the specific gravity rising from 1.005 at a depth of 46 m to 1.10 at the bottom. It is interesting to speculate on the origin of the salt concentration gradient in the lake. Wilson⁵ suggested that at some period in the past the climate was such that the Onyx River did not supply appreciable water to Lake Vanda. In these conditions the lake level would have dropped until only a few feet of concentrated calcium chloride remained. When the climate changed, the Onyx would

flow during the summer and fresh water would have flowed on top of this strong salt solution. Since then the salts would have been diffusing upwards. If such a model is assumed, it is possible to calculate the time when this climatic change occurred. On the above assumptions it is possible to calculate that Lake Vanda rose relatively suddenly 1,200 yr ago⁵. This same approach has been used to calculate diffusion ages for Lake Bonney^{6,7}.

When snow falls on a snow field or land surface it contains small quantities of salts, the principal cations being sodium and calcium and the principal anion being chloride with smaller amounts of Mg^{2+} , K^+ and SO_4^{2-} and other materials. The chemical composition of atmospheric precipitation has been the subject of much study (see refs 8–10). Although the chemicals in precipitation are generally believed to come from the ocean, the actual composition is not merely diluted seawater but has been modified by separation processes at the ocean surface. The actual composition varies as one moves away from the sea due, it is believed, to the existence of two types of aerosol particles in the atmosphere. The first type is large, falls out rapidly and has the composition of bulk seawater. Particles of the second type are very much smaller and have a composition very different from seawater. They have, for example, higher K^+/Na^+ and $\text{Ca}^{2+}/\text{Mg}^{2+}$ ratios than bulk seawater. Both types of aerosols arise from the bursting of bubbles on the surface of the ocean^{8,9}. In order to explain the distribution of salts in Antarctic soils and the composition of Antarctic lakes, the following hypothesis is presented.

Discussion

Over very long periods of time small quantities of snow fall on the land surface and on sublimation leave salts behind. These salts include Na^+ , K^+ , Ca^{2+} , Mg^{2+} , SO_4^{2-} , Cl^- . It is proposed that fluctuations in relative humidity, due to the occasional snowfall and the summer-to-winter change, cause some of the more deliquescent salts to slowly percolate downwards through the soil. The more deliquescent will reach the frozen water table. At high altitudes, for example, on the top of the Asgard and Olympus Mountains and in the Alatna Valley, only the least deliquescent, for example sodium sulphate, can remain in the soil; the remaining salts, CaCl_2 , MgCl_2 and much of NaCl , eventually reach the frozen water table where during the late summer they move down slope along the surface of the frozen water table. In the more arid situations at lower altitudes, the NaCl (77% equilibrium relative humidity at 0 °C, ref. 11) cannot reach the frozen water table and on the sides of the McMurdo dry valleys; for example, above Don Juan Pond and at similar altitudes along the sides of the Wright Valley, large quantities of NaCl can be found in the 'soils' and under boulders. At higher altitudes in the same region, almost pure Na_2SO_4 and $\text{Na}_2\text{SO}_4 \cdot 2\text{H}_2\text{O}$ can be found in the same position (equilibrium relative humidity of sodium sulphate = 98% at 0 °C, ref. 11). The calcium chloride (equilibrium relative humidity = 44% at 0 °C, ref. 11) and magnesium chloride (equilibrium relative humidity = 38% at 0 °C, ref. 11) eventually move to the lowest part of the drainage system—in the case of the Wright Valley this is Lake Vanda and Don Juan Pond. Thus it is proposed that we have huge geochemical separation systems in the Wright Valley and other Antarctic dry areas, the salt fractionation being controlled by relative humidity and the deliquescent properties of the possible salts that can form from the anions and cations left by evaporated snow. Very saline groundwater can be detected flowing along the surface of the frozen water table in the late summer. For example, groundwater flowing along the top of the frozen water table in the south fork of the Wright Valley in late January 1974 had the following composition: 61,400 p.p.m. Cl^- , 11 p.p.m. SO_4^{2-} , 18,100 p.p.m. Ca^{2+} , 7,000 p.p.m. Mg^{2+} , 9,900 p.p.m. Na^+ , 635 p.p.m. K^+ (sample collected by A.T.W. and analysed by A. Field). The groundwater is high in Ca^{2+} with respect to Na^+ , and low in SO_4^{2-} with respect to Cl^- and has the chemical composition necessary to explain the origin of the calcium chloride type lakes discussed above.

In the above discussion, relative deliquescence is proposed as the primary mechanism for the separation of water soluble salts in the Antarctic dry areas. For completeness, freezing must also be considered. Eutectic temperature would at first seem to provide an alternative parameter that could regulate the composition of the liquid phase, and such a mechanism will probably play some part in the initial separations. Evidence that relative deliquescence is the dominating mechanism is provided by the 'calcium chloride' type lakes and the 'groundwater' (see data above) which are not enriched in K^+ with respect to Na^+ . This would be expected on the relative deliquescent model proposed (equilibrium relative humidity of 0 °C of NaCl = 77% and KCl = 88%¹¹) whereas on an eutectic temperature model the reverse would be expected. When brines of the general composition expected in Antarctic soils are frozen¹², the K^+/Na^+ ratio remains constant until -8.2 °C, at which point sodium sulphate decahydrate forms. The K^+/Na^+ ratio continues to increase with falling temperatures until -22.9 °C at which point sodium chloride dihydrate forms. The ratio further increases until at -36 °C, sylvite (KCl) separates as does the dodecahydrate of magnesium chloride. On an eutectic freezing model, one would have expected potassium to follow magnesium and be enriched in the 'calcium chloride' type lakes and ponds. In fact, the reverse is true, leading one to conclude that it is the relative deliquescent mechanism rather than a freezing or a solubility mechanism which controls the separation of soluble salts in polar deserts.

The 'groundwater' described above can also flow into the sides of glaciers and explains the origins of the clear brown saline 'amber ice' often found in the basal layers of the local glaciers in the Antarctic dry areas (see, ref. 13). It also provides an explanation for the fact that some glacial meltwater streams have a higher $\text{Ca}^{2+}/\text{Na}^+$ ratio than the snow falling on their neves.

All the lakes in the Antarctic dry areas in practice have salts entering from both groundwater flow and meltwater from snow and glacier ice. Even glacial meltwater streams seem to have a significant contribution from groundwater salt, particularly early in the season. The fact that the composition of inflow streams varies throughout the summer and does not match the composition of the lake surface waters is an anomaly which is readily explained when the importance of the salt contribution by groundwater is realised. The interfingering of high Ca^{2+} groundwater to its appropriate density level can produce layers of very different $\text{Ca}^{2+}/\text{Na}^+$ in a lake. See, for example, the data of Bell¹⁴ for Lake Miers. Layers of high $\text{Ca}^{2+}/\text{Na}^+$ ratio presumably contributed by groundwater can be seen at 17 m below the surface, and at the bottom.

Samples collected from the bottoms of Lake Bonney (east lobe) and Lake Vanda contain 14.9 ± 1.1 and 6.4 ± 0.4 tritium units respectively. Surface snow in this area contains about 100 tritium units. This was 'discovered' when I was attempting to provide the New Zealand Institute of Nuclear Science Tritium Dating Laboratory with tritium-free water for calibration purposes. It is generally believed that the bottom of the east lobe of Lake Bonney has been out of contact with the atmosphere for at least 60 yr⁷ and the bottom of Lake Vanda for 1,200 yr⁵. As the half life of tritium is 12.3 yr, there should be essentially no tritium present in the bottom waters of either lake. I confirmed that the 'age' of the bottom waters of Lake Vanda is at least 1,000 yr by radiocarbon measurements. If the brine flowing along the top of the frozen water table enters a density-stratified lake it will interfinger to its appropriate density layer; and I propose that this is the mechanism whereby tritium has entered an apparently isolated geochemical system. Note that the inflow of saline groundwater into Antarctic lakes as proposed here would introduce small amounts of carbonate which had been in at least partial equilibrium with the atmosphere. This introduces complications for those attempting to use radiocarbon dating techniques to determine the times at which the lake levels, and hence climate, changed. The interfingering of saline groundwater into Antarctic lakes also provides an additional

mechanism for the formation of chemical stratification over and above that proposed by Wilson⁵. The use of diffusion calculations to determine the times at which lake levels changed must be made with caution with particular attention being paid to how all the ionic ratios vary with depth.

The rather startling implication of this hypothesis is that much of the million tons of chloride contained in Lake Vanda has entered from groundwater flow. The amounts of Na⁺ and SO₄²⁻ needed to bring the overall composition back to that of snow water are to be found in the soils on the lower sides of the valley (as NaCl) and high up on the Asgard and Olympus Ranges (as Na₂SO₄ and Na₂SO₄·2H₂O). Such a system would have taken a long time to develop and implies a great age to the Wright Valley geochemical system. A conclusion supported on other evidence¹⁵.

Once the importance of saline groundwater is accepted one can appreciate how certain materials such as the strontium of Jones and Faure¹ can be leached from the glacial debris over which it flows. Thus, the isotopic ratio of the strontium in the sodium chloride type lakes could be of marine origin, whereas the calcium chloride type lakes where the salts are largely of groundwater origin could have components derived from the leaching of rock—which does seem to be the case¹.

The geochemical separation process described here can explain most of the apparently conflicting observations made during geochemical studies in the Antarctic dry areas and the unusual geochemistry of the area. Such geochemical separation systems have developed in Antarctica and not in other arid parts of the world because the Antarctic dry areas are in a truly

rainless desert. The infrequent but violent rains which occur even in the most arid of the non-polar deserts would destroy such a salt separation system. Although the geochemical situation described here is unusual on the Earth as a whole, being restricted to the rainless polar deserts, it may be expected to be found in other parts of the Solar System. For example, one could expect there to be soluble salts on the surface of Mars; and as Mars is also a rainless desert, one could expect similar geochemical separations to those described here to have taken place.

Much of the work reported here was made possible by grants from the New Zealand University Grants Committee and the cooperation of the New Zealand Antarctic Division, DSIR, the US Navy and US Antarctic Research Program.

Received 15 December, accepted 16 May 1979.

1. Jones, L. M. & Faure, G. *Earth planet. Sci. Lett.* **3**, 101–106 (1967).
2. Charlesworth, J. K. *The Quaternary Era* Vol. 1 (Arnold, London, 1957).
3. Torji, T. & Ossaka, J. *Science* **149**, 975–977 (1965).
4. House, D. A. *et al. J. Chem. Educ.* **43**, 502–505 (1966).
5. Wilson, A. T. *Nature* **201**, 176–177 (1964).
6. Shirlcliffe, T. G. L. *J. geophys. Res.* **69**, 5257–5268 (1964).
7. Hendy, C. H., Wilson, A. T., Popplewell, K. B. & House, D. A. *N.Z. J. Geol. Geophys.* **20**, 1103–1122 (1977).
8. Wilson, A. T. *Nature* **184**, 99–101 (1959).
9. Wilson, A. T. *Nature* **186**, 705–706 (1960).
10. Junge, C. E. *Air Chem. Radioactivity, Int. Geophys. Ser. 4*, (Academic, New York, 1963).
11. Washburn, E. W. (ed.) *International Critical Tables of Numerical Data*, Vol. 3 (McGraw-Hill, New York, 1928).
12. Thompson, T. G. & Nelson, K. H. *Am. J. Sci.* **254**, 227–238 (1956).
13. Holdsworth, G., thesis, Ohio State Univ. (1969).
14. Bell, R. A. I. *N.Z. J. Geol. Geophys.* **10**, 540–556 (1967).
15. Wilson, A. T. in *Palaeoecology of Africa, the Surrounding Islands and Antarctica VIII* (ed. van Zinderen Bakker, E. M.) 23–25 (Balkema, Cape Town, 1973).

Embryonic development of identified neurones: differentiation from neuroblast to neurone

Corey S. Goodman* & Nicholas C. Spitzer

Department of Biology B-022, University of California at San Diego, La Jolla, California 92093

Individually identified neurones are sufficiently large and accessible in grasshopper embryos to permit visualisation and impalement with intracellular microelectrodes from the time of their birth to their maturation. In this article part of the temporal pattern of differentiation from an identified neuroblast (precursor cell) to a group of identified neurones is described.

THE development of nerve cells involves the acquisition of specific neuronal phenotypes. Precursor cells proliferate and give rise to progeny which differentiate into neurones. During neuronal differentiation, processes grow out, arborise and reach their targets where synaptic connections are formed. The neurones acquire a responsiveness to neurotransmitters, an ability to make action potentials and a capacity for their own neurotransmitter synthesis, storage and release. The temporal sequence in which these and other phenotypes appear has not yet been determined for any group of nerve cells *in vivo*. A knowledge of the timing in any one instance would be a useful step in understanding the rules and mechanisms of neuronal development. For example, must the appearance of one specific phenotype precede the appearance of another?

It is possible to identify single unique neurones in the nervous systems of both invertebrates^{1–3} and vertebrates⁴; they are

usually identified on the basis of their characteristic location, morphology, physiology and biochemistry. Identified neurones are attractive for the study of neuronal development, as it is possible to follow the timetable of single cells from their birth to their maturation, thus avoiding the problems involved in examining the developmental average of a heterogeneous and asynchronous population.

We have been studying the embryonic central nervous system (CNS) of the grasshopper and have focused our attention on the development of the dorsal unpaired median (DUM) neurones because they represent a population of relatively large, identifiable and accessible embryonic cells. The adult CNS of the grasshopper consists of the brain and a chain of segmental ganglia running along the ventral surface of the thorax and abdomen. The DUM cells are a group of identified neurones (80–100) situated on the dorsal surface of the thoracic and abdominal ganglia, near the midline. The properties of the mature DUM neurones are well known from studies of adults^{5–11}. The embryonic DUM neurones can be visualised and impaled with intracellular microelectrodes at the earliest stages of their development. Furthermore, the lineage of these neurones can be determined by serial examination of dissected preparations of increasing age, with interference contrast optics, as has been demonstrated for some neurones in the nematode^{12,13}. We are now able to write down part of the developmental timetable of some of the identified DUM neurones of the thoracic ganglia in the embryo of the grasshopper. Preliminary accounts of this work have appeared^{14,15}.

The present study begins with the day 5 embryo, a stage at which the precursor cells (neuroblasts) for the segmental ganglia

*Present address: Department of Biological Sciences, Stanford University, Stanford, California 94305.

have already differentiated from the epidermis. The embryonic nervous system of the day 8 embryo is shown in Fig. 1A. At the centre of each body segment are the left and right plates of ventral neuroblasts (Fig. 1Aiii) which give rise to the chain of segmental ganglia¹⁶⁻²⁰.

Each ventral neuroblast is $\sim 30\ \mu\text{m}$ in diameter (Fig. 1Aiv). There are about 31 ventral neuroblasts in each left or right plate. Bate²⁰ found that the pattern and number of neuroblasts at early embryonic stages was at least as constant as that of the identified neurones of the adult segmental ganglia. Each neuroblast (NB) is a stem cell which will not in itself become a neurone. Rather, each NB maintains its large size while budding off a chain of smaller cells (Fig. 1Av). Each of these progeny divides into two cells which are postmitotic and differentiate into neurones. The neuroblasts eventually degenerate at precise times during embryogenesis²⁰.

Each segmental array of neuroblasts also contains a single unpaired neuroblast, called the median or dorsal unpaired median (DUM) neuroblast, at the posterior end of the segment and dorsal to the paired plates of ventral neuroblasts^{16,18,20} (Fig. 1Aiii, iv, v). As Wheeler¹⁶ first described, the progeny of the single DUM neuroblast extend anteriorly across the dorsal surface of the developing ganglion, quite separate from the neuronal progeny of the ventral neuroblasts. Thus, neuronal differentiation in this lineage can be viewed by removing the embryo from the egg, removing the chain of segmental ganglia from the embryo, desheathing the dorsal surface of the ganglion, and viewing the dorsal surface under a $\times 40$ water immersion lens of a compound microscope with Nomarski interference contrast optics.

The dorsal unpaired median neurones

In the adult grasshopper the somata (cell bodies) of most of the 3,000 neurones in each thoracic ganglion (pro-, meso- and metathoracic) lie on the ventral surface or lateral edge of the ganglion and the neuropil lies dorsal to them. These cells are paired, that is, each cell has a contralateral homologue that is its mirror image. The left and right plates of ventral neuroblasts give rise to these ventral paired neurones^{16,18,20}. However, there

is a group of 80–100 somata clustered along the midline on the dorsal surface of the ganglion. These neurones, called dorsal unpaired median neurones⁷, have the following distinctive characteristics. The cells thus far examined in the adult are unpaired and have medium neurites which bifurcate into bilaterally symmetrical processes. The DUM neurones are the only neurones in the segmental ganglia whose somata generate overshooting action potentials. The soma spikes of the DUM neurones depend on both Na^+ and Ca^{2+} while the axon spikes are predominantly Na^+ dependent¹¹.

Whereas DUM neurones have certain characteristics in common, many if not all of them can be uniquely identified by their morphology. The largest DUM neurone in the grasshopper metathoracic ganglion (diameter $\sim 40\text{--}60\ \mu\text{m}$) is a unique identified neurone with axons emerging bilaterally in nerve 5 and innervating the extensor tibiae (ETi) muscle in the femur of the leg (this neurone is called DUMETi (ref. 7)).

Only 6–8 of the DUM neurones have large-diameter somata ($40\text{--}60\ \mu\text{m}$); most are somewhat smaller ($15\text{--}30\ \mu\text{m}$ in diameter). The somata of the DUM neurones stain with neutral red^{8,9}, a dye which selectively stains monoamine-containing neurones in the leech²¹ and lobster^{22,23}. Both the soma and axons of DUMETi have been shown to contain octopamine⁹. In fact, most if not all of the DUM neurones (particularly the large ones) contain octopamine²⁴.

Electrical and dye coupling

All cell types examined at early embryonic stages of the grasshopper are electrically coupled to each other; development involves a process of selective uncoupling. Dissected embryos were perfused with a simple saline (NaCl 125 mM; KCl 3 mM; CaCl_2 10 mM; HEPES buffer 5 mM; adjusted to pH 7.4). Cells were impaled with intracellular microelectrodes, each with recording and current passing capability. Neuroblasts ($\sim 30\ \mu\text{m}$ diameter) had resting potentials of -60 to $-80\ \text{mV}$; non-neuronal cells ($\sim 10\ \mu\text{m}$ diameter) often had recorded resting potentials of -40 to $-60\ \text{mV}$, that may have been the result of damage during impalement. Records made by inserting two microelectrodes into the same ventral or dorsal neuroblast at

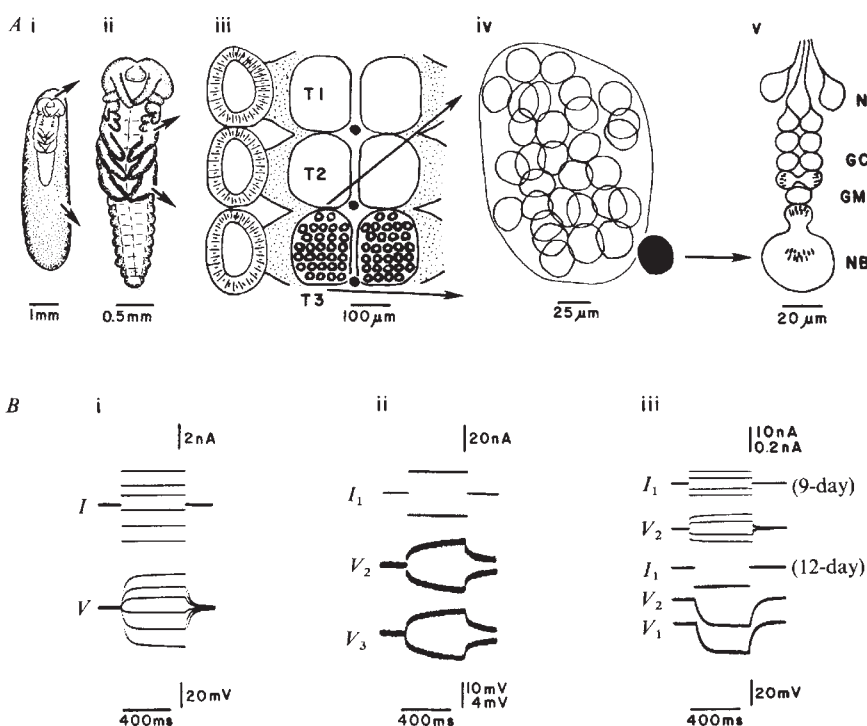


Fig. 1 A, Grasshopper embryo (*Schistocerca nitens*) at day 8; diagrams and camera lucida drawings of living specimens. Location of neuroblasts (precursor cells) and formation of neurones are shown (hatching occurs on day 20 at 35°C). (i) Embryo at day 8 viewed within the egg case. The embryo metabolises yolk, shown stippled. (ii) The day 8 embryo viewed from the ventral surface, showing the three prominent thoracic limb buds and the equivalent segmental appendages forming the mouth parts. (iii) Removal of the thoracic limb buds exposes the paired plates of ventral neuroblasts (enlarged ectodermal cells, open circles), and the single dorsal unpaired median (DUM) neuroblast (filled circles), also called the median (MNB) neuroblast, at the posterior margin of each segmental array of ventral neuroblasts. The contiguous chain of paired plates of neuroblasts gives rise to the chain of segmental ganglia (T1, pro; T2, meso; T3, metathorax). (iv) Camera lucida drawing of the left hemiganglionic plate of neuroblasts in the metathoracic segment, also showing the dorsal unpaired median neuroblast, viewed with a $\times 40$ water immersion lens and interference contrast optics. (v) A schematic drawing showing a neuroblast (in this case, the DUM neuroblast) as a stem cell; it maintains its large size while budding off a chain of ganglion mother cells (GMC) which each divide into two ganglion cells (GC) and then differentiate into neurones (N). B, Input resistance (R_{in}) and extent of electrical coupling of cells at different stages of development of the grasshopper embryo. (i) R_{in} for a ventral neuroblast at day 8 of development ($\sim 10\ \text{M}\Omega$). Current (I) and voltage (V) electrodes in the same cell. (ii) Electrical coupling of one ventral neuroblast (I_1) to a second ventral neuroblast (V_2) and to an epithelial cell of the limb bud (V_3), in a day 8 embryo. (iii) Electrical coupling of DUM neuroblast (I_1) to nearby progeny cell (V_2) at day 9 (note large current); DUM neuroblast is still coupled to the rest of the embryo at this stage. Coupling of DUM neuroblast (I_1) to nearby progeny cell (V_2) at day 12 (note small current); DUM neuroblast becomes uncoupled from most of the rest of the embryo at day 10. R_{in} for the DUM neuroblast (V_1) and for the coupled network is $\sim 200\ \text{M}\Omega$.

coupling of DUM neuroblast (I_1) to nearby progeny cell (V_2) at day 9 (note large current); DUM neuroblast is still coupled to the rest of the embryo at this stage. Coupling of DUM neuroblast (I_1) to nearby progeny cell (V_2) at day 12 (note small current); DUM neuroblast becomes uncoupled from most of the rest of the embryo at day 10. R_{in} for the DUM neuroblast (V_1) and for the coupled network is $\sim 200\ \text{M}\Omega$.

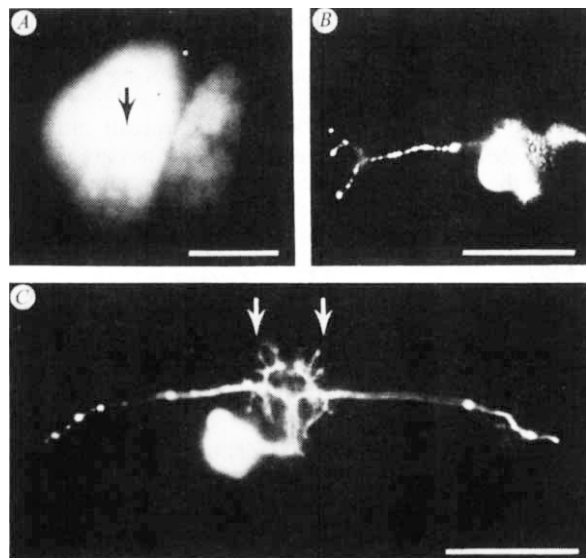


Fig. 2 Lucifer Yellow (450 MW) injections of a neuroblast and two identified neurones of grasshopper embryos. Fluorescence viewed with UV illumination. Scale bar, 100 μ m. **A**, Injection of a ventral neuroblast (arrow) in the left hemiganglionic plate of neuroblasts of the metathoracic segment of an 8-day embryo. Dye spreads rapidly through other ventral neuroblasts and some of their progeny in the left plate, and within a few minutes through cells in the right plate. Note absence of spread to adjacent limb buds. **B**, **C**, Injection of identified neurones in the metathoracic ganglion of a 13-day embryo. Cell bodies appear unusually large because of intense fluorescence. The dye does not spread to other cells. Axons are truncated at points of exit from the ganglion. **B**, DUM 3, 4, 5: median neurite bifurcates, and left axon sends branches out nerves 3, 4 and 5 (top to bottom). **C**, DUMETi (DUM 5): axons extend out nerve 5 only. Arrows indicate beginnings of relatively symmetrical central arborisations.

easily stages (Fig. 1Bi) show that the input resistance (R_{in}) is quite low (~ 10 M Ω), probably as a result of the extensive electrical coupling. The input resistance of a single ventral neuroblast was found to be about 150 M Ω after mechanically removing it from the network of neighbouring cells to which it was coupled, and impaling it with two electrodes. This is likely to be less than the true value because of possible damage during the isolation.

Ventral neuroblasts are electrically coupled to one another within a hemiganglion (Fig. 1Bii). Coupling also exists between left and right hemiganglia, between the DUM neuroblast and ventral neuroblasts, and between neuroblasts of a given segment and those of a segment either anterior or posterior to it. Furthermore, the neuroblasts are coupled to the smaller and morphologically distinct epithelial cells (termed cap cells) which lie just over them, and even to the more distant epithelial cells of the developing limb bud (Fig. 1Bii). Confirmation of the positions of the electrode tips and the identities of the coupled cells was achieved by injecting cells with the dyes Fast Green or Procion Rubine (2% dye in 0.25 M KCl). No evidence for rectification of electrical coupling was seen. In all such experiments, the coupling recorded between any two cells disappeared when either intracellular electrode was withdrawn from the cell. Ventral neuroblasts become electrically uncoupled from the limb buds at day 10, but remain electrically coupled to each other until they stop dividing and degenerate by day 15.

We examined the movement of low molecular weight tracers from the interior of one cell to the interior of another by injecting the dye Lucifer Yellow (450 MW), which has a high fluorescent quantum efficiency and rapid rate of diffusion. When iontophoresed into single cells (from electrodes filled with 5% dye in distilled H_2O), it will often spread to other cells to which the injected cell is electrically coupled²⁵. When a single ventral neuroblast is filled with Lucifer Yellow in a day 8 embryo, the dye spreads to other ventral neuroblasts and some of their progeny in the hemiganglion within a few minutes (Fig. 2A). The fluorescence can be detected in the contralateral hemiganglion

shortly thereafter, but spreads to anterior and posterior ganglia much more slowly. However, dye was not seen to spread to limb-bud epithelium, even though the neuroblasts are electrically coupled to the limb-bud cells. In control experiments, dye was not taken up by the cells when simply applied in the bath.

We next examined the electrical and dye coupling of the DUM neuroblast and its progeny. The DUM neuroblast begins dividing at day 6. At first, all of its progeny are electrically coupled to the DUM neuroblast (Fig. 1Biii). Until day 9, large currents injected into the DUM neuroblast produce small voltage changes in progeny cells (Fig. 1Biii, top), consistent with the extensive electrical coupling of the DUM neuroblast to the rest of the embryo. As early as day 10 (Fig. 1Biii, bottom, similar records from day 12) much smaller currents produce large voltage changes in progeny cells. Day 10 is the time at which the DUM neuroblast becomes electrically uncoupled from most embryonic cells other than its own progeny. Input resistance (R_{in}) for the coupled network of neuroblast and progeny (as well as the R_{in} of the DUM neuroblast) increases from < 10 M Ω at day 9 to ~ 200 M Ω by day 12 (Fig. 1Biii). All progeny remain electrically coupled to the neuroblast until day 11, when the oldest progeny (identified DUM neurones described below) are first becoming electrically uncoupled. (Mature DUM neurones are not electrically coupled to one another.)

Lucifer Yellow injected into the DUM neuroblast at day 7 spreads throughout all of its progeny. However, at day 8, its

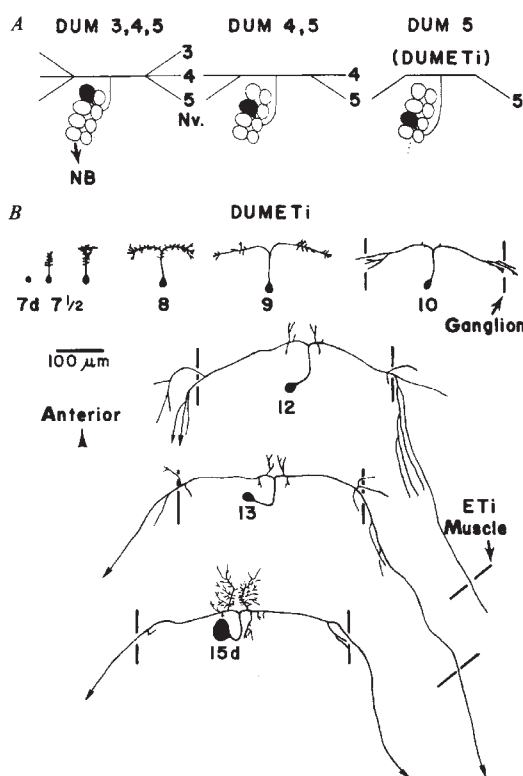


Fig. 3 **A**, DUM neuroblast gives rise to identified DUM neurones. The DUM neuroblast (NB), not shown here, gives rise to chains of progeny; some of the oldest of these progeny become identified DUM neurones. At days 12–13, we find two chains of progeny, generally two cells in width. We have examined the identities of the first three cells on the outside of the oldest chain of progeny ($n = 10$). The morphology of the DUM neurones is revealed by intracellular injection of Lucifer Yellow. In this schematic diagram, the most anterior cell on the outside bifurcates and bilaterally extends out peripheral nerves 3, 4, and 5; the second cell bilaterally extends out nerves 4 and 5; and the third cell extends out bilaterally only nerve 5 (this cell is DUMETi). **B**, Morphological development of the single identified DUM neurone, DUMETi ($n \geq 3$ for each time point). The DUM neuroblast begins dividing at about day 6. Many extra branches and growth cones are produced at the distal end, whether it be in the central nervous system or periphery, and these extra neurites later disappear as the correct branch for DUMETi continues to grow. 'Ganglion' marks the border of the metathoracic ganglion, and 'ETi muscle' marks the proximal boundary of the extensor tibiae muscle in the femur of the limb bud. Arrows mark points of axonal extension beyond the margins of the figure.

oldest progeny are dye-uncoupled. Dye does not spread into or out of these identified DUM neurones, although they are still electrically coupled to other cells (Fig. 2B, C).

DUM neuroblast gives rise to identified DUM neurones

The evidence that the DUM neuroblast gives rise to identified DUM neurones is described by considering day 12–13 embryos, although these cells can be followed from day 6 by serial examination of dissected preparations of increasing age. Viewing the desheathed dorsal surface of the metathoracic ganglion reveals chains of progeny extending anteriorly from the DUM neuroblast. Cells close to the neuroblast are young progeny which are the result of recent divisions (the neuroblast continues dividing until it degenerates on day 16), whereas those at the anterior (distal) end of the elongating chains are the oldest progeny. The DUM neuroblast is $\sim 30\ \mu\text{m}$ in diameter, the progeny at first are $\sim 10\ \mu\text{m}$ and quickly enlarge to $\sim 15\ \mu\text{m}$ in diameter. Several of the somata increase to a diameter of over $40\ \mu\text{m}$ during embryogenesis. Some of the oldest progeny of the DUM neuroblast become identified as DUM neurones because within a few days they assume the characteristic location (dorsal midline), the characteristic morphology (median neurites which bifurcate and extend bilaterally, Figs 2, 3), the characteristic physiology (spiking somata, Fig. 5) and the characteristic histochemistry (somata which stain with Neutral Red²⁴).

At days 12–13, we see two chains of progeny, each being generally two cells in width, with one chain extending anteriorly on the left side and the other chain on the right side of the midline. Initially the older chain is dorsal to the younger (day 7). Between days 7 and 11, the plane in which the two chains lie changes in orientation, from dorso-ventral to horizontal. When the orientation changes, there is a 50% probability ($n = 24$) that the dorsal (older) chain will assume either the left or right position in the horizontal plane. It can easily be distinguished, however, by the morphology of its cells throughout this period (C. M. Bate, C.S.G. and N.C.S., in preparation).

We have examined the identities of the first three cells on the outside of the oldest chain of progeny, and find the same identified neurone in the same position along the chain ($n = 10$) (Fig. 3A). The morphology of the DUM neurones is revealed by intracellular injection of Lucifer Yellow, and later visualisation in either living or fixed tissue, with UV illumination (Fig. 2B, C). (Tissue is fixed in 4% formaldehyde, dehydrated in an ethanol series, cleared in methyl benzoate and mounted in fluoromount.) The most anterior cell on the outside bifurcates and extends bilaterally out as peripheral nerves 3, 4 and 5; the second cell on the outside bifurcates and extends bilaterally out as peripheral nerves 4 and 5; and the third cell on the outside bifurcates and extends bilaterally out only peripheral nerve 5. The axons of this third cell run out nerve 5 to the extensor tibiae (ETi) muscle; this cell is DUMETi. In the mesothoracic ganglion, the first cell also bifurcates and extends bilaterally out as peripheral nerves 3, 4 and 5. Thus, from animal to animal, the cell in the same position along the chain, presumably having the same mitotic ancestry, develops into the same identified neurone.

Morphological development of identified DUM neurones

The outgrowth of processes by the single identified DUM neurone, DUMETi, was followed as a function of developmental age ($n \geq 3$ for each time point). Axonal outgrowth of DUMETi begins on about day 7 with the extension of a median neurite (Fig. 3B). At this time, the framework for the neuropil, nerve connectives and peripheral nerves already exists in the form of two longitudinal fibre tracts, two commissures in each segment, and lateral fibre tracts to the periphery. The distal end of the growing axon produces multiple branches with growth cones extending in different directions, often for over $20\ \mu\text{m}$ in the wrong direction. As the correct branch continues to extend

and send out branches at its distal tip, the extra incorrect neurites disappear, leaving a relatively naked proximal axon.

By day 8, the axon of DUMETi has bifurcated and both lateral axons appear fuzzy because of the presence of many fine branches. During day 9, the distal end of the axon grows out to the edge of the ganglion, and many of the proximal branches disappear. By day 10, the neurone sends axons out many of the available peripheral nerves and branches of those nerves, even though ultimately these extra branches will disappear and the axon of DUMETi will extend out only nerve 5 (and branch 5b1d) to the ETi muscle. By day 12, the distal axon reaches the ETi muscle. Three prominent features at this age are the occurrence of extra peripheral branches, the start of the central bilateral arborisations and the relatively naked central axon. By day 13, the axon has extended over the entire ETi muscle, but the cell body and arborisations are still relatively small.

After the axon of DUMETi reaches its target, there are three clear morphological changes: the extra peripheral neurites disappear, leaving behind only one axon in nerve 5; the central arborisations greatly increase in the extent and number of their

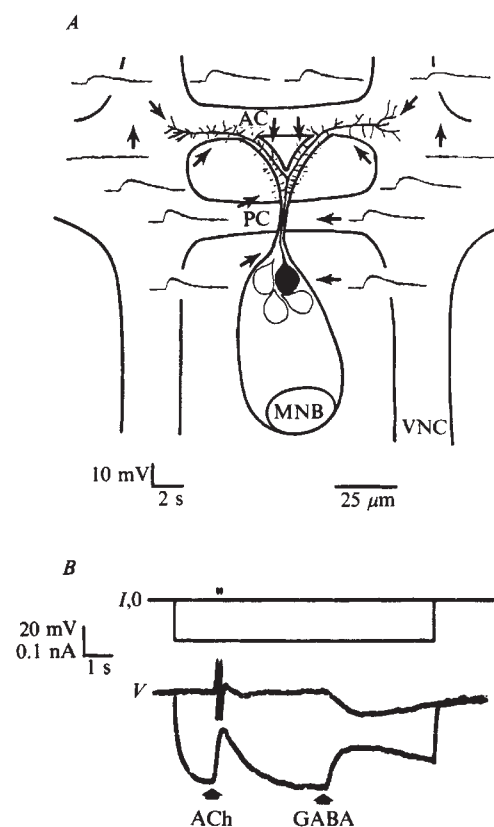


Fig. 4 Chemosensitivity of identified DUM neurones in the grasshopper embryo (the three oldest progeny as described above). **A**, Response to iontophoretic application of ACh in a day 8 embryo. The diagram shows the embryonic neuropil of a single segmental ganglion (T3), as viewed from the dorsal surface (anterior at top). The DUM neuroblast (NB) is shown with its packet of progeny; those most anterior (oldest) send their axons anteriorly in a median bundle of processes which cross the posterior commissure (PC) and bifurcate near the anterior commissure (AC). The extent of branching of an individual neurone was determined by intracellular injection of Lucifer Yellow. The longitudinal connectives extend into the ventral nerve cords (VNC) and are seen on either side. The lateral neurite extends (beyond the margin of the figure) into three fibre tracts which become peripheral nerves 3, 4 and 5. The response to application of ACh at the points indicated by the arrows were recorded by an intracellular electrode in the cell body. Processes are sensitive to ACh over their whole length. Small vertical or lateral displacements of the iontophoretic electrode abolished the response. **B**, Responses to sequential iontophoretic application of ACh and GABA from a double-barrelled micropipette, recorded with an intracellular electrode in a day 13 embryo. At a resting potential of $-55\ \text{mV}$, ACh depolarises the cell to threshold, eliciting overshooting action potentials, while GABA hyperpolarises the cell. When the cell is hyperpolarised by injected current, the response to ACh is larger, but remains subthreshold; the cell is then depolarised by GABA.

fine branches; and the cell body greatly enlarges. A similar initial overproduction of distal branches and eventual disappearance of these extra branches was found also for DUM 3, 4, 5 and DUM 4, 5.

Onset of chemosensitivity

At day 7, neither the DUM neuroblast nor its progeny show responses to bath application of acetylcholine (ACh) or γ -aminobutyric acid (GABA) at 10^{-3} M. In contrast, by 13 days the oldest identified DUM neurones (described above) are sensitive to bath application and iontophoresis of both of these agents. This latter result is illustrated in Fig. 4B, which shows the intracellularly recorded responses of a developing DUM neurone to iontophoretic application of ACh and GABA from a double-barrelled electrode. These responses involve a membrane conductance increase (not illustrated). The observed reversal potential for GABA is -70 mV, while the extrapolated reversal potential for ACh lies above $+20$ mV. The ACh response is abolished by removal of Na^+ (replacement with choline); the GABA response is abolished by removal of Cl^- (replacement with isethionate). The ACh response is reduced by the nicotinic antagonists curare (10^{-4} M), hexamethonium (10^{-3} M) and decamethonium (10^{-3} M); nicotine is an agonist. However, α -bungarotoxin does not significantly reduce the response at 10^{-6} M. The ACh response is unaffected by the muscarinic antagonists atropine (10^{-6} M) and quinuclidinyl benzilate (QNB, 10^{-7} M); pilocarpine has no effect (10^{-3} M). The GABA response is rapidly blocked by picrotoxin (10^{-4} M); muscimol is an agonist.

The oldest DUM neurones first become sensitive to ACh and GABA at day 8 of development, and seem to become sensitive to both neurotransmitters at the same time. Young progeny in close proximity to the DUM neuroblast show no response to iontophoretic application of ACh. In contrast, the oldest progeny at the distal end of the chain, which are still electrically coupled (Fig. 1Biii) and already possess neurites (Fig. 3B), are depolarised by iontophoretic application of ACh onto their somata. (No desensitisation is seen in response to prolonged application of these agents.)

The processes of the oldest DUM neurone are also depolarised by iontophoretic application of ACh on day 8. Processes seem to be sensitive over their entire length. It seems that chemosensitivity is distributed over the whole surface of the cell, at or soon after the time that it first appears. Identified DUM neurones thus develop chemosensitivity to ACh and GABA while still electrically coupled, and may be sensitive to other neurotransmitters as well. The reversal potential and ionic dependence of the response to ACh and GABA appear similar at days 8, 13 and 18 of development.

Onset of electrical excitability

We have examined the onset of electrical excitability in the oldest DUM progeny by observing the responses of cells to depolarising current pulses injected through an intracellular recording electrode. They are generally incapable of producing action potentials at day 11. They become electrically excitable between days 11 and 12 (Fig. 5A), while they are becoming electrically uncoupled. It seemed possible that voltage-dependent channels were not revealed while the cells were electrically coupled because of the long time constant of the coupled network of cells, preventing depolarisation to threshold before channel inactivation occurred. This possibility was tested by exposing the cells to veratridine 10^{-5} g ml $^{-1}$, which specifically depolarises cells possessing voltage-dependent Na^+ channels²⁶. The treatment has no effect on the resting potentials of any of the ventral neuroblasts, the DUM neuroblast, or its progeny up to day 11. In contrast, the oldest DUM progeny are depolarised by the treatment by day 12 (we thank A. Willard for suggesting this approach).

Between days 11 and 12, the oldest DUM neurones first produce small action potentials which fail to overshoot the zero

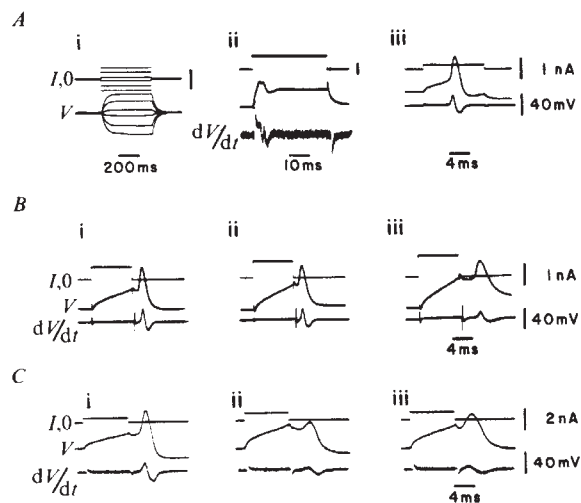


Fig. 5 Electrical excitability of the oldest DUM neurones in the grasshopper embryo. **A**, Responses in the oldest DUM progeny between days 11 and 12. (i) The DUM neuroblast is impaled with one microelectrode to inject current, and an electrically coupled progeny is impaled with a second microelectrode to record voltage. These cells have resting potentials of -50 to -70 mV and appear electrically inexcitable. Depolarisation and hyperpolarisation of the neuroblast and its progeny by 50 mV reveals a constant membrane resistance. (ii) As the cells become uncoupled they develop neurite action potentials in which the inward current is carried predominantly by Na^+ (data not shown). The early embryonic neurites bifurcate at a T-junction, and in many cases the first embryonic axon spikes have the adult characteristic of appearing double-peaked; each peak in the adult and presumably in the embryo corresponds to a separate site of initiation along one of the two lateral axons. (iii) Current injected into a DUM neurone at a later time yields an overshooting action potential that appears to arise in the cell body. **B**, Ionic dependence of the somatic action potential in a day 12 embryo. The overshoot of the spike seen in normal saline (i) is not greatly affected by replacement of Na^+ with Tris (ii) or addition of 30 mM Co^{2+} (iii). **C**, Ionic dependence of the action potential of a DUM neurone in a day 18 embryo. The overshoot (i) is reduced by either removal of Na^+ (ii) or addition of Co^{2+} (iii). In both day 12 and day 18 action potentials, simultaneous removal of Na^+ and addition of Co^{2+} (not shown here) abolishes the action potential.

potential and often have a characteristic double-peaked appearance (Fig. 5A). These action potentials are likely to be produced in the neurites, because this response has been shown in adult DUM neurones to be the sum of action potentials arising from two symmetrical neurites distal to the point of bifurcation¹⁰. These action potentials depend largely on an inward Na^+ current, because they are rapidly blocked by removal of Na^+ or addition of 10^{-9} g ml $^{-1}$ tetrodotoxin (TTX). These cells and those in older embryos are depolarised by 25 mV or more when exposed to veratridine. This depolarisation is blocked by prior removal of Na^+ or addition of TTX.

Within a day, the oldest DUM neurones are capable of producing overshooting action potentials, 2 – 4 ms in duration, that appear to arise in the cell body ($R_{in} \approx 300$ M Ω). These action potentials depend on Na^+ and Ca^{2+} for their inward current. Removal of Na^+ or addition of 30 mM Co^{2+} (to block Ca^{2+} channels) has little effect on the overshoot (Fig. 5B), although simultaneous application of both treatments abolishes the action potential. However, by 18 days, blockage of either inward current reduces the overshoot in these same neurones (Fig. 5C). It is known that in at least one adult DUM neurone, DUMETi, the blockage of either inward current alone is sufficient to abolish the action potential¹¹. The basis for these developmental changes is not yet known.

Development of later phenotypes

The transmitter of the DUM neurones, octopamine, begins to appear in the thoracic ganglion on day 13, after the axons of the oldest DUM neurones reach their peripheral target²⁴. Shortly thereafter, the somata of the oldest DUM neurones begin to stain with Neutral Red.

The first spontaneous synaptic activity, from unidentified synaptic input, is recorded in DUM neurones at 15 days, 1 week after the appearance of chemosensitivity. The source of these synaptic inputs is not yet known.

Temporal sequence of development

There is a distinct order in the acquisition of particular phenotypes by the identified DUM neurones that are the oldest progeny of the DUM neuroblast (Fig. 6). This developmental sequence occurs first in the oldest progeny; younger progeny may be as much as several days behind in their differentiation.

Some of the phenotypes initially express change during the course of further embryonic development whereas others do not. There are changes in the morphology of identified neurones: the initial fine branches and the extra peripheral neurites all eventually disappear. Similarly an initial overproduction and later disappearance of many central branches around the site of the main arborisation has been observed in an identified neurone in the brain of the grasshopper (D. Bentley and A. Raymond, personal communication). In other systems, gap junctions²⁷ and profiles of chemical synapses²⁸ appear transiently on cell bodies shortly before the outgrowth of their processes. In the embryonic DUM cells, there appears to be no change in the ionic basis of the responses to ACh and GABA, as indicated by the ion substitution experiments and the reversal potential determinations. In contrast, there is a change in the ionic requirements for electrical excitability of the cell body, consistent with, although not qualitatively as great as, the changes reported for the development of excitability in other neurones²⁹⁻³³.

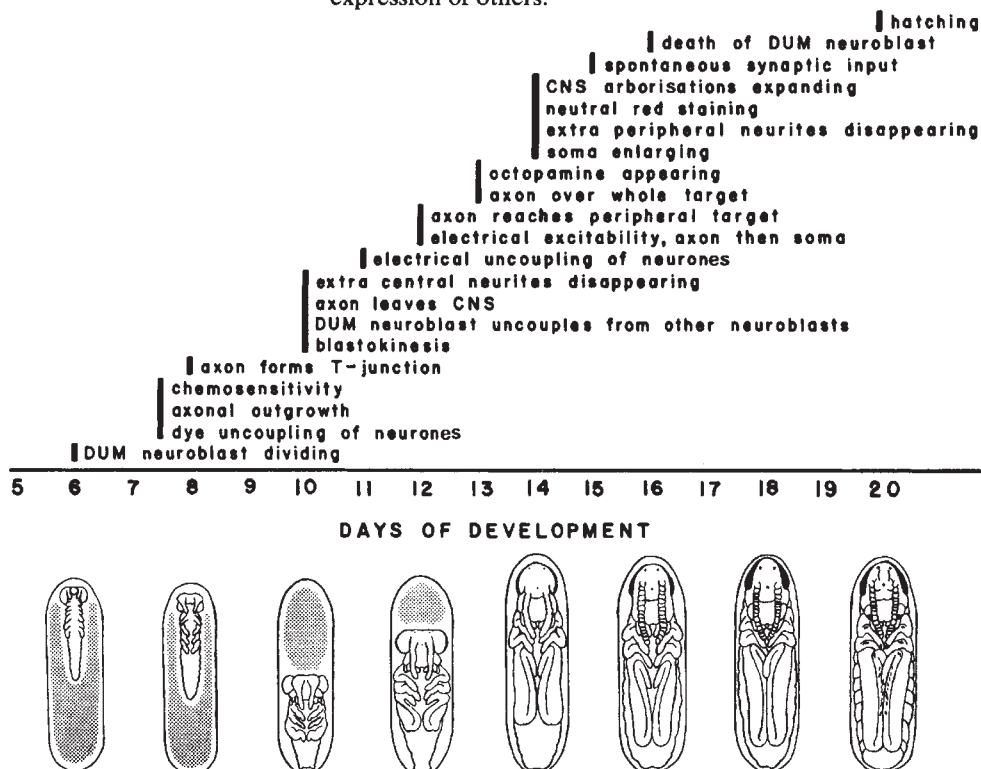
The initially extensive electrical coupling and eventual uncoupling are consistent with observations in other embryos³⁴⁻³⁶. The outgrowth of processes and the appearance of chemosensitivity of the identified DUM neurones begin after their terminal mitoses but while they are still extensively electrically coupled. The observation that cells become dye uncoupled before they become electrically uncoupled suggests that there may be some progressive restriction on the molecular size of substances that can pass freely between cell interiors. Further work will be needed to determine the mechanisms involved in these changes and the role of uncoupling in further differentiation.

How general is the sequence of appearance of neuronal phenotypes observed in the DUM neurones? Studies of the development of amphibian spinal cord have shown that Rohon-Beard neurones become electrically excitable a substantial time after their final DNA synthesis^{30,37}. They are already depolarised by GABA at this time, but it is not known how early chemosensitivity first appears³⁸. The development of electrical excitability of the neurites of embryonic amphibian neurones seems to precede the cell body by a small time interval³³. The identity of the neurotransmitter of Rohon-Beard cells is not known. The known portion of the developmental sequence for these cells is compatible with the sequence for DUM neurones.

Neurotransmitter synthesis is a relatively early phenotype for vertebrate superior cervical ganglion cells^{39,40}, whereas it is a relatively late phenotype for invertebrate DUM neurones. However, vertebrate ciliary ganglion neurones show a late large increase in neurotransmitter synthetic enzyme activity after their axons have reached and synapsed on their targets⁴¹. Neurotransmitter synthesis may differentiate earlier in other neurones in grasshopper embryos, since high levels of ACh are already present by day 5, before differentiation of central neurones²⁴. Sanes and Hildebrand⁴² found that antennal sensory cells in the pupae of the moth *Manduca* were capable of synthesising and storing ACh before the axons arrived at their targets in the central nervous system. If the onset of neurotransmitter synthesis requires particular environmental cues^{43,44}, then these cues may exert their influence at different points in the lives of these different types of cells. Alternatively, different types of neurones may simply be reading out different intrinsic programmes.

Further work will be needed before any causal relationship in the sequence of phenotypes of the embryonic DUM neurones can be established or ruled out. However, it is worth noting that several events are closely linked temporally. The outgrowth of processes and the onset of chemosensitivity occur at about the same time. Cells become electrically uncoupled and electrically excitable at about the same time. Neurotransmitter storage is detected just after cell processes have reached their targets. Cell body enlargement, Neutral Red staining, the disappearance of extra peripheral neurites, and the expansion of central arborisations occur at this time. It may now be possible to manipulate one phenotype and examine the consequences for the later expression of others.

Fig. 6 The developmental timetable for the oldest DUM neurones of the grasshopper embryo *in vivo* [DUM 3, 4, 5; DUM 4, 5; and DUM 5 (DUMETi)]. The egg is laid at day zero, and at 35 °C hatching occurs at day 20. The time of onset or first appearance of a phenotype in these neurones is indicated by a vertical bar. The morphology of the embryo, viewed from the ventral surface in the egg case, is illustrated below as traced from camera lucida projections. Note that the embryo makes a major rotation (blastokinesis) at day 10, and thus from day 10 onward the embryo is viewed from the opposite side of the egg case.



We thank Michael Bate and Michael O'Shea for stimulating discussions, Janet Lamborghini and Darwin Berg for critical reading of the manuscript, William Kristan and Allen Selverston for providing facilities, James Coulombe for technical assistance, and Nino Cooley for preparation of the figures. Bungarotoxin was supplied by P. Ravdin and D. Berg, Lucifer Yellow by Walter Stewart, and QNB by M. Nirenberg and Hoffman-LaRoche. This work was supported by NIH and NSF grants to N.C.S. and a Helen Hay Whitney Fellowship to C.S.G.

Received 27 December 1978; accepted 4 June 1979.

1. Kuffler, S. W. & Nicholls, J. G. *From Neuron to Brain* (Sinauer, Boston, 1976).
2. Kandel, E. R. *Cellular Basis of Behavior* (Freeman, San Francisco, 1976).
3. Hoyle, G. (ed.) *Identified Neurons and Behavior of Arthropods* (Plenum, New York, 1977).
4. Faber, D. S. & Korn, H. (eds) *Neurobiology of the Mauthner Cell* (Raven, New York, 1978).
5. Plotnikova, S. I. *J. evolut. biochem. Physiol.* **5**, 339–341 (1969).
6. Crossman, A. R., Kerkut, G. A., Pitman, R. M. & Walker, R. J. *comp. Biochem. Physiol.* **40A**, 579–596 (1971).
7. Hoyle, G., Dagan, D., Moberly, B. & Colquhoun, W. J. *exp. Zool.* **187**, 159–165 (1974).
8. O'Shea, M. & Evans, P. *Soc. Neurosci.* **3**, 187 (1977).
9. Evans, P. & O'Shea, M. *Nature* **270**, 275–279 (1977); *J. exp. Biol.* **73**, 235–260 (1978).
10. Heitler, W. J. & Goodman, C. S. *J. exp. Biol.* **76**, 63–84 (1978).
11. Goodman, C. S. & Heitler, W. J. *Soc. Neurosci.* **3**, 426 (1977); *J. exp. Biol.* (in the press).
12. Sulston, J. E. *Phil. Trans. R. Soc. B* **275**, 287–297 (1976).
13. Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **56**, 110–156 (1977).
14. Goodman, C. S. & Spitzer, N. C. *Soc. Neurosci.* **4**, 113 (1978).
15. Spitzer, N. C. & Goodman, C. S. *Soc. Neurosci.* **4**, 127 (1978).
16. Wheeler, W. M. *J. Morph.* **4**, 337–343 (1891); **8**, 1–160 (1893).
17. Baden, V. J. *Morph.* **60**, 159–190 (1936).
18. Roonwal, M. L. *Phil. Trans. R. Soc. B* **227**, 175–244 (1937).

19. Anderson, D. T. in *Developmental Systems: Insects* (eds Counce, S. J. & Waddington, C. H.) (Academic, New York, 1972).
20. Bate, C. M. *J. Embryol. exp. Morph.* **35**, 107–123 (1976).
21. Stuart, A. E., Hudspeth, A. J. & Hall, Z. W. *Cell Tissue Res.* **153**, 55–61 (1974).
22. Wallace, B. G., Talamo, B. R., Evans, P. D. & Kravitz, E. A. *Brain Res.* **74**, 349–355 (1974).
23. Evans, P. D., Kravitz, E. A., Talamo, B. R. & Wallace, B. G. *J. Physiol., Lond.* **262**, 51–70 (1976).
24. Goodman, C. S., O'Shea, M., McCaman, R. E. & Spitzer, N. C. *Science* **204**, 1219–1222 (1979).
25. Stewart, W. W. *Cell* **14**, 741–759 (1978).
26. Narahashi, T. *Physiol. Rev.* **54**, 813–889 (1974).
27. Lopresti, V., Macagno, E. R. & Levinthal, C. *Proc. natn. Acad. Sci. U.S.A.* **70**, 433–473 (1974).
28. Schacher, S. & Kandel, E. R. *Soc. Neurosci.* **4**, 126 (1978).
29. Spitzer, N. C. & Baccaglini, P. I. *Brain Res.* **107**, 610–616 (1976).
30. Baccaglini, P. I. & Spitzer, N. C. *J. Physiol., Lond.* **271**, 93–117 (1977).
31. Baccaglini, P. I. *J. Physiol., Lond.* **283**, 585–604 (1978).
32. Spitzer, N. C. & Lamborghini, J. E. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1641–1645 (1976).
33. Willard, A. L. *Soc. Neurosci.* **3**, 123 (1977); and in preparation.
34. Potter, D. D., Furshpan, E. J. & Lennox, E. S. *Proc. natn. Acad. Sci. U.S.A.* **55**, 328–336 (1966).
35. Warner, A. E. in *Simple Nervous Systems* (eds Newth, D. R. & Usherwood, P. N. R.) (Arnold, London, 1975).
36. Blackshaw, S. E. & Warner, A. E. *J. Physiol., Lond.* **255**, 209–230 (1976).
37. Spitzer, N. C. & Spitzer, J. L. *Am. Zool.* **15**, 781 (1975).
38. Spitzer, N. C. *Soc. Neurosci.* **2**, 204 (1976).
39. Cochar, P., Goldstein, M. & Black, I. B. *Soc. Neurosci.* **4**, 109 (1978); *Proc. natn. Acad. Sci. U.S.A.* **75**, 2986–2990 (1978).
40. Rothman, T. P., Gershon, M. D. & Holtzer, H. *Dev. Biol.* **65**, 322–341 (1978).
41. Chiappinelli, V., Giacobini, E., Pilar, G. & Uchimura, H. K. *Physiol., Lond.* **257**, 749–766 (1976).
42. Sanes, J. R. & Hildebrand, J. G. *Dev. Biol.* **52**, 105–120 (1976).
43. LeDouran, N. M., Renaud, D., Teillet, M. A. & LeDouran, G. H. *Proc. natn. Acad. Sci. U.S.A.* **72**, 728–732 (1975).
44. LeDouran, N. M., Teillet, M. A., Ziller, C. & Smith, J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2030–2034 (1978).
45. Bentley, D., Keshishian, H., Shankland, M. & Raymond, A. J. *Embryol. exp. Morph.* (in the press).

letters

Accretion on massive black holes in galactic nuclei

STELLAR collisions and/or tidal break-up of stars by a massive black hole^{1–3} accompanied by subsequent accretion of the released gas onto the hole play a crucial part in most black hole models of quasars and active galactic nuclei. It is usually assumed that an accretion disk forms around the hole due to the large orbital momentum of a disrupting star. However, we show here that the accretion mostly has disk characteristics only when $M < M_{\text{cr}}$ and becomes quasi-spherical when $M > M_{\text{cr}}$.

The appearance of the critical mass M_{cr} is determined by the fact that the reservoir of the accreting matter is replenished discretely. Moreover, the time scale between two stellar break-ups τ_d (determined by the consumption rate) does not generally coincide with the evolution time scale, τ_{evol} , of the first star remnant. If $\tau_{\text{evol}} < \tau_d$, the accretion of this remnant ceases before the next star can be disrupted. Hence an accretion disk will form producing a rather discrete picture in time. Whenever $\tau_{\text{evol}} \gg \tau_d$ many other stars will be disrupted during the evolution of the first star's disk debris, producing remnants orientated randomly in space. Collisions between these remnants will lead to the more or less spherical shape of the resulting gas cloud, and its accretion onto the black hole will proceed in this nearly spherical region.

On considering quantitative details we first obtain the evolution time scale τ_{evol} for the disk remnant of a disrupted star (called 'disk' here). After tidal break-up of a first star (which may be accompanied by its thermal flare⁴ the gas remnant expands and after one or few revolutions around the hole, that is the time scale $\tau_{\text{form}} \approx 3 \times 10^2 (M_{\text{h}}/10^6 M_{\odot})$ yr forms an elliptical disk with axis depending on both energy and angular momentum of the orbiting star. If they are conserved after the disruption of the

star, the minimal and maximum distances of the disk boundary are related by

$$\frac{r_{\text{min}}}{r_{\text{max}}} = \frac{a(1-e)}{a(1+e)} \approx \frac{1}{2}(1-e) \quad (1)$$

Here $e \approx 1$ is the eccentricity of the disk; r_{min} is of the order of the tidal radius $R_t = 2.2 r_*(M_{\text{h}}/m_*)^{1/3} \approx 1.5 \times 10^{13} M_6^{1/3}$ cm for a black hole of mass $M_{\text{h}} = 10^6 M_6 M_{\odot}$; $r_{\text{max}} = (4/3) r_* M_{\text{h}}/m_* \approx 10^{17} M_6$ cm; $m_* \approx 1 M_{\odot}$ and $r_* \approx 1 R_{\odot}$ are the typical stellar mass and radius. Therefore, $1-e^2 = 6 \times 10^{-4} M_6^{-2/3}$, and the semi-axis of the disk are

$$a = \frac{1}{2} r_{\text{max}} \approx 5 \times 10^{16} M_6 \text{ cm} \\ b = a(1-e^2)^{1/2} \approx 1.2 \times 10^{15} M_6^{2/3} \text{ cm} \quad (2)$$

Viscous friction causes the disk to accrete gradually on to the hole. This accretion proceeds in the optically thick manner. The viscosity in the disk is determined mainly by turbulent transfer processes, so we may write $\nu = v_i l$, where v_i is the velocity of turbulent motions, and l is the free path length. The value of l is approximately equal to the disk's half-thickness $H = \nu_s/v_{\phi}$ (v_s is the sound velocity and $v_{\phi} = (GM/r)^{1/2}$ is the keplerian velocity). Assuming $v_i \approx v_s$, one obtains $\nu \approx \nu_s^2/v_{\phi}$ which by a factor ~ 1 coincides with the Ichimaru's theory of magnetohydrodynamic turbulence in the accreting disk⁵:

$$\nu = \left(\frac{2}{\pi}\right)^{3/2} \left(\frac{kT}{m}\right)^{1/2} H \\ H = 1.18 r \left(\frac{GM}{r}\right)^{-1/2} \left(\frac{2kT}{m}\right)^{1/2} \quad (3)$$

(m is the ion mass). The rate of viscous dissipation per unit

volume is $\epsilon_v = \frac{1}{2} \nu \rho r^2 (\partial \Omega / \partial r)^2 \sim \Sigma (kT/m)^{1/2} \Omega^2 (\rho, \Sigma = 2H\rho, \text{ and } \Omega \text{ being volume density, surface density, and angular velocity respectively}). The maximum rate of viscous dissipation is reached in the outer domain of the disk near the radius$

$$R_m = 1.2 \times 10^{13} M_6^{1/3} \dot{M}_{-6}^{2/3} \text{ cm} \quad (4)$$

($\dot{M}_{-6} = \dot{M}/10^{-6} M_\odot \text{ yr}^{-1}$) from when the bound-free and free-free absorption dominate the electron scattering opacity. In this region ($r \geq R_m$) the balance between the rate of viscous dissipation and the rate of radiative loss yields the temperature of plasma in the disk

$$T = 4.4 \times 10^4 M_6^{1/4} \dot{M}_{-6}^{3/10} r_{13}^{-3/4} \text{ K} \quad (5)$$

From equations (3) and (5) we find the viscosity coefficient

$$\nu = 8.5 \times 10^{15} M_6^{-1/4} \dot{M}_{-6}^{3/10} r_{13}^{3/4} \text{ cm}^2 \text{ s}^{-1} \quad (6)$$

Knowing the main parameters of the steady state disk (including the dependence of ν on r) enables us to simulate the secular evolution of the viscous disk following the model of Lynden-Bell and Pringle⁶. The mass of the disk diminishes as it accretes on to the black hole according to the law $m = m_i(1 + t/\tau_{\text{evol}})^{-2/5}$ where $m_i \approx 1M_\odot$ is the initial disk mass, R_i is its characteristic size, and

$$\begin{aligned} \tau_{\text{evol}} &= \frac{32 R_m^{3/4} R_i^{5/4}}{75 \nu(R_m)} \\ &\approx 5.4 \times 10^5 M_6^{1.84} \text{ yr} \end{aligned} \quad (7)$$

is the evolution time scale. In equation (7) we use $R_i = (ab)^{1/2}$ to account for the elliptical disk form and adopt $\dot{M} = m/\tau_{\text{evol}}$ in equation (4). A more approximate, but essentially the same estimate for

$$\tau_{\text{evol}} \approx \Sigma / \dot{\Sigma} \approx \Omega \left| \frac{\partial \Omega}{\partial r} \right|^{-1} \frac{r}{\nu} \Big|_{r=R_i} \approx 1.8 \times 10^6 M_6^{1.84} \text{ yr} \quad (8)$$

is obtained from time-dependent equations for disk accretion.

Now, comparing τ_{evol} from equation (7) with the time scale τ_d between successive tidal disruptions of two stars¹

$$\tau_d \approx 5 \times 10^2 M_6^{-4/3} v_{200} n_7^{-1} \text{ yr} \quad (9)$$

we obtain for the critical black hole mass

$$M_{\text{cr}} \approx 1.1 \times 10^5 n_7^{-0.32} v_{200}^{0.32} M_\odot \quad (10)$$

For conditions similar to those in the nucleus of our Galaxy (star density $n \approx 10^7 \text{ pc}^{-3}$, velocity dispersion $v = 200 \text{ km s}^{-1}$) we find $M_{\text{cr}} \sim 10^5 M_\odot$, and for a dense globular cluster ($n = 5 \times 10^4 \text{ pc}^{-3}$, $v = 20 \text{ km s}^{-1}$) $M_{\text{cr}} \sim 3 \times 10^5 M_\odot$ is obtained.

At $M < M_{\text{cr}}$ the disk evolves faster than the next star enters the Roche lobe of the black hole. Therefore when the black hole mass is small enough, the accretion takes place in the disk regime and ceases before the next star breaks up. As follows from equation (7), at $t \leq \tau_{\text{evol}}$ the accretion rate exceeds Eddington critical rate $\dot{M}_{\text{Edd}} \approx 2 \times 10^{-2} M_6 (M_\odot \text{ yr}^{-1})$ if the black hole mass $M_h < 4 \times 10^4 M_\odot$. For smaller masses the fact that the accretion rate $\dot{M}$ is proportional to $t^{-7/5}$ is very critical during the time interval $t_s \approx 7M_4^{-2.03} \tau_{\text{evol}} = 0.8 \times 10^3 M_4^{-0.19} \text{ yr}$. However, the accretion keeps its disk character as the luminosity of the hole $L \leq (0.6 - 1.0) L_{\text{Edd}}$ (ref. 7). Mass outflow from the disk is possible during t_s , but the value of t_s is negligible compared with τ_d (equation (9)) if $M_4 \leq 1$. As for the smaller black hole masses, we believe that at $M_h < 10^4 M_\odot$ the tidal disruption of stars is not the main fuel for black holes in galactic nuclei.

When $M = M_{\text{cr}}$, the characteristic time scales τ_{evol} and τ_d are of the same order of magnitude. In this case and especially when $M > M_{\text{cr}}$, star-disk collisions shorten the time scale τ_d . The accumulation of star remnants around the black hole leads to the spherisation of the resulting cloud and to the subsequent quasi-spherical accretion. So when $M \gg M_{\text{cr}}$, we may assume that the accretion will be quasi-steady with the rate equal to that of stellar consumptions, analogous to the case of accretion on to a compact object in close binary systems where the accretion rate is controlled by matter outflow rate from the normal star.

We have assumed, in accordance with most models, that the accreting gas is fuelled by tidally disrupted stars. However, there are three other possible sources of this gas: (1) gas loss by stars during their evolution; (2) star collisions; (3) accretion of intergalactic gas.

As for the source (1), the present rate of gas loss by stars in the galactic nuclei is of the order of $M^{-1} \dot{M} \sim 10^{-12} - 10^{-11} \text{ yr}^{-1}$ (refs 8, 9). This gas probably has a large angular momentum and forms a disk which supplies the black hole with gas at a rate less than $10^{-4} - 10^{-5} M_\odot \text{ yr}^{-1}$ when the total mass of stars in the nucleus is $10^7 - 10^8 M_\odot$. The real accretion rate seems to be much smaller as part of gas transforms into young stars, and another part is expelled from the nucleus by supernovae explosions. (As an example, the nucleus of our Galaxy shows evidence both of increased number of young stars and of supernova remnants.) It is easily seen that at $M_h < 10^4 - 10^5 M_\odot$, gas lost by stars may become the main fuel for the black hole, but the accretion remains in a disk regime.

In contrast, when the black hole mass is large enough ($M_h > 2 \times 10^7 n_7^{1/2} r_c^{3/4} M_\odot$, the core radius, r_c , (pc)) the breakup of stars by collisions begins to dominate the tidal disruptions. In this case the evolution time for a disk formed (which may be estimated analogous to equation (7)) seems to be much larger than the time interval between successive collisions, so that the accretion remains quasi-spherical.

Only when the main source of the accreting gas is the flow of intergalactic gas which has a very large angular momentum relative to the galactic centre, is the accretion expected to have a disk character irrespective of the black hole mass. However, the intergalactic reservoir of gas may, in principle, be distinguished from the intragalactic reservoir as in nuclei of galaxies belonging to a close pair or group, the activity is expected to be simultaneous (coincidence scheme)¹⁰. It is possible that some point radio sources in the nuclei of elliptical and spiral galaxies which are members of the same groups or pairs^{11,12} show evidence for the accretion of intergalactic gas or flow of outlying clouds.

V. G. GURZADYAN
L. M. OZERNOY

P. N. Lebedev Physical Institute,
USSR Academy of Sciences,
Moscow, USSR

Received 23 January; accepted 10 May 1979.

1. Hills, J. G. *Nature* **254**, 295-298 (1975).
2. Rees, M. *Phys. Scripta* **17**, 193-200 (1978).
3. Ozerney, L. M. & Reinhardt, M. *Astrophys. Space Sci.* **59**, 171-192 (1978).
4. Lydskij, V. V. & Ozerney, L. M. *Pis'ma Astr. Zh.* **5**, 28-33 (1979).
5. Ichimaru, S. *Astrophys. J.* **214**, 840-855 (1977).
6. Lynden-Bell, D. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **168**, 603-637 (1974).
7. Bisnovatyi-Kogan, G. S. & Blinnikov, S. I. *Astr. Astrophys.* **59**, 111-125 (1977).
8. Gislér, G. R. *Astr. Astrophys.* **51**, 137-150 (1976).
9. Shields, G. A. & Wheeler, J. C. *Astrophys. J.* **222**, 667-674 (1978).
10. Ozerney, L. M. *Astr. Zh.* **47**, 265-269 (1970).
11. Condon, J. J. & Dressel, L. L. *Astrophys. J.* **221**, 456-467 (1978).
12. Kerr, F. J. in *IAU. Symp.* No. 77, 224-225 (1978).

Observation of large 2.2 μm polarisation in 3C345

INFRARED polarisation observations of the quasars 3C345 and 3C273 are reported here. The highly polarised and variable optical emission of BL Lacertae objects is correlated with similar high polarisation and variability in the IR, and is clearly the product of non-thermal processes. IR polarisation by dust, or by a combination of dust and non-thermal processes, is also found in Seyfert galaxies¹⁻⁵. These observations suggest that the measurement of IR polarisation in quasars be used to investigate the often dominant IR emission in these sources. However, so far only one measurement of polarisation at wavelengths as long as 2.2 μm in a quasar, 3C273, has been reported⁴.

A clear distinction exists between the variable, steep slope QSOs which often show strong and variable optical polarisation, and the less variable QSOs with lower spectral index which do not^{6,7}. The former resemble the BL Lac objects in the steepness of their optical spectra and in their optical polarisation. 3C345 is such a quasar with variable visual and IR flux⁸ and highly variable visual polarisation^{9,10}. 3C273 may vary somewhat on a long time scale, but this slow (if present at all) variation and the small optical and IR polarisation suggest that 3C273 is a member of the latter group. As these two quasars seem to show different phenomena, we have observed them in order to compare their IR polarisation.

Linear polarisation at 2.2 μm was measured with a chopper-coupled polarimeter (to be described elsewhere¹¹) at the Kitt Peak National Observatory 4-m telescope. The sensitivity to polarisation is a function of both the intensity and polarisation of the sources; as a rough guide, in one hour of integration time 1% polarisation can be measured on a source with K magnitude of 11 at the 4-m telescope. Instrumental polarisations were measured to be 0.42% (J, 1.25 μm), 0.38% (H, 1.6 μm) and 0.12% (K, 2.2 μm). The data were corrected for this effect by subtraction of the appropriate Stokes parameters. A further correction was made for the polariser efficiency determined by measurement of the transmission of crossed polarisers.

The position angle was calibrated by observing the polarisation of VI Cyg 12 and computing the difference between the measured angle, as determined by the arbitrarily positioned polariser, and the angle at K of $\theta = 113^\circ$ determined by Dyck and Jones¹². The K polarisation that we found for VI Cyg 12, $P_K = 1.39 \pm 0.04\%$, is in reasonable agreement with the Dyck and Jones measurement of $P_K = 1.22 \pm 0.12\%$.

We first observed 3C345 on 5 May 1977 (Table 1). Very large 2.2 μm polarisation of $20.6 \pm 7.2\%$ was noticed at that time, but because the signal-to-noise ratio did not give a completely convincing result, we felt that the measurement required corroboration. This became possible in April 1978 (Table 1), and this time 3C345 was observed on three nights. The 2.2 μm polarisation was again very large. The magnitudes of the polarisations are close to those found in the previous year and the orientations also seems similar. However, this agreement may well be coincidental as the IR polarisation may be variable as is the visual. Further observations are required to establish IR variability.

Remarkably, the IR polarisation of 3C345 is much greater than the visual. The visual measurement that was made on 25 April (at the same time as the two IR measurements of that night) was noisy because of sky background but established an upper limit to the B (0.44 μm) polarisation of 6%. The measurements of 26 and 27 April are consistent with this result. They also show that despite being much smaller in magnitude, the visual polarisation had approximately the same orientation as the IR on the previous nights.

In extragalactic sources in which polarisation has been measured, the IR polarisation is either equal to (for example, BL Lac)¹ or less than (for example, MK421, NGC4151)² the visual polarisation. High IR polarisation accompanied by much smaller visual polarisation is as yet a phenomenon unique to 3C345 amongst extragalactic objects.

Table 1 Polarisation measurements of 3C345 and 3C273

Object	Date	λ (μm)	$P \pm \delta P$	$\theta \pm \delta\theta$
3C345	5 May 1977	2.2	20.6 ± 7.2	32 ± 10
	23 April 1978	2.2	23.8 ± 6.4	67.0 ± 7.5
	24 April 1978	2.2	21.9 ± 4.2	65.4 ± 5.5
	25 April 1978	0.44 (B)	< 6	—
		1.6	9.7 ± 6.1	—
		2.2	32.6 ± 6.5	65.0 ± 5.6
	26 April 1978	0.44 (B)	4.5 ± 0.7	68 ± 9
	27 April 1978	0.44 (B)	2.2 ± 0.7	43 ± 19
3C273	5 May 1977	2.2	0.07 ± 0.23	—

We have observed 3C273 intermittently since 1975, and observed it again in April 1978. In April 1978 we observed the source with the new instrument and obtained a much lower limit on the polarisation than in our earlier studies. The formal measurement (Table 1) of $P_K = 0.07 \pm 0.23\%$ corresponds to an upper limit (2σ) of 0.46%. This is consistent with the possible detection of $P_K = 0.33 \pm 0.14\%$ of Kemp *et al.*⁴. Thus 3C273 does not seem to be an object with active IR polarisation.

Optical polarisation of 3C345 has been studied for more than a decade. These data and the radio polarisation have been reviewed by Kinman¹⁰. Between 1967 and 1974 the optical polarisation varied between 1% and 12% and was accompanied by position angle fluctuations of 20° – 110° . Kinman suggests that the position angle shows some regularity, lying between 80° and 110° during 'outbursts' of high flux, and this direction may be related to the position angle of super-relativistic expansion, $105^\circ \pm 5^\circ$. Stockman *et al.* have also suggested an alignment of the optical and radio components¹³. During our observations, however, the apparently parallel visual and IR polarisation directions were not aligned with the expansion direction.

The spectrum and variability of 3C345, and now the polarisation, suggest that the IR emission is non-thermal. Polarisation by scattering off dust seems unlikely to give the observed wavelength dependence. It would probably be difficult for any scattering mechanism to account for the large IR polarisation unless there is very great asymmetry. The synchrotron mechanism can give large visual and IR polarisation, although the peculiar wavelength dependence remains to be explained.

Any explanation for the wavelength dependence of the IR polarisation based on Faraday rotation is unlikely for several reasons. First, in optically thick sources, the polarisation tends to decrease towards longer wavelengths, opposite to the sense found in 3C345 (ref. 14). Second, in this mechanism the 2.2 μm and visual polarisation parallelism would be unaccounted for. And third, Faraday rotation at IR wavelengths would require larger fields and densities than are generally encountered in similar sources.

It may be more plausible to look for a dilution mechanism because the approximate parallelism of the IR and visual polarisations suggest that the IR and optical sources are related. However, spectra of 3C345 taken by Wampler¹⁵ do not show emission lines in the B passband strong enough to dilute continuum polarisation. Wampler also showed that the absence of a Balmer break indicates that thermal continuum emission is unlikely to be strong in 3C345.

An explanation may lie in the models of optical polarisation of Kinman *et al.*⁹. They suggest that three sources contribute to the polarisation. Two of these, one a slowly and one a rapidly varying component, are assumed to be polarised by approximately 17%, but in unrelated directions. Then as intensity variations occur, the sum of the polarisations combines to give the observed 2–11% variations. The IR polarisation that we observed is close to Kinman *et al.*'s value of 17% for the components. The observed wavelength dependence of polarisation would result if, for example, the slowly varying component extends from the visual to the IR, and the rapidly varying component, perhaps with flatter slope, interferes with the visual, but not the IR polarisation. Some support for a multicomponent model is found from the observed dual component structure in the radio. The agreement of the two measurements separated by a year implies that the IR polarised component is only slowly varying.

The high degree of IR polarisation places 3C345 among the 'violent variables' shown by Nordsieck¹⁶ to occupy the upper right portion of a plot of polarisation against spectral index. The polarisation requires a high degree of order in the magnetic field. From Nordsieck's equations we find his parameter δ describing the 'squash' of the magnetic field (where $P(H, \chi) = P_I [H_x^2 / (1 + \delta^2) + H_y^2 / (1 - \delta^2)]^{1/2}$ is the probability that an element of volume has a field H perpendicular to the line of sight in direction χ) is $\delta \approx 0.5$ (with spectral index 1.2). Similarly the ratio of a directed field to an isotropic field (Nordsieck's case IIa)

is $\varepsilon/H \approx 0.4$. Evidently considerable long-range order exists in the source.

The detection of strong IR polarisation in a quasistellar source presents another point of similarity between quasars and BL Lac objects. IR polarisation measurements can evidently provide insights into the structure and emission mechanisms of quasistellar objects, especially the optically active ones which often also have steep slopes. IR and visual polarisation activity seems to be correlated in quasars as it is in BL Lac objects. However, as an example with greater polarisation in the IR than in the visible has now been found, it may be useful not to preselect these objects by optical polarisation criteria alone; non-thermal sources that may be partially obscured by other visual emission may appear in the IR.

We thank Jeffrey Moller for many contributions to this project, including the computer interfacing and software, and N. Jones-Knacke for helpful comments. This work was supported by a grant from the University Awards Committee of the SUNY Research Foundation. R.F.K. and R.W.C. are guest observers at the Kitt Peak National Observatory.

R. F. KNACKE

*Astronomy Program,
Department of Earth and Space Sciences,
State University of New York,
Stony Brook, New York 11794*

R. W. CAPPS

*Institute for Astronomy,
University of Hawaii,
Honolulu, Hawaii 96822*

M. JOHNS

*McGraw-Hill Observatory,
Sells Star Route,
Tucson, Arizona 85735*

Received 26 February; accepted 24 May 1979.

- Knacke, R. F., Capps, R. W. & Johns, M. *Astrophys. J. Lett.* **210**, L69–L70 (1976).
- Rieke, G. H., Lebofsky, M. J., Kemp, J. C., Coyne, G. V. & Tapia, S. *Astrophys. J. Lett.* **218**, L37–L42 (1977).
- Knacke, R. F. & Capps, R. W. *Astrophys. J. Lett.* **192**, L19–L22 (1974).
- Kemp, J. C., Rieke, G. H., Lebofsky, M. J. & Coyne, G. V. *Astrophys. J. Lett.* **215**, L107–L110 (1977).
- Lebofsky, M. J., Rieke, G. H. & Kemp, J. C. *Astrophys. J.* **222**, 95–99 (1978).
- Visvanathan, N. *Astrophys. J.* **179**, 1–20 (1973).
- Kinman, T. D. *Astrophys. J.* **205**, 1–5 (1976).
- Oke, J. B., Neugebauer, G. & Becklin, E. E. *Astrophys. J.* **159**, 341–355 (1970).
- Kinman, T. D., Lamla, E., Ciurla, T., Harlan, E. & Wirtanen, C. A. *Astrophys. J.* **152**, 357–376 (1968).
- Kinman, T. D. *Nature* **267**, 798–800 (1977).
- Capps, R. W. & Knacke, R. F. (in preparation).
- Dyck, H. M. & Jones, T. J. *Astr. J.* **83**, 594–597 (1978).
- Stockman, H. S., Angel, J. R. P. & Miley, G. K. *Astrophys. J. Lett.* **227**, L55–L58 (1979).
- Gardner, F. F. & Whiteoak, J. B. A. *Rev. Astr. Astrophys.* **4**, 245–292 (1966).
- Wampler, E. J. *Astrophys. J.* **147**, 1–12 (1967).
- Nordsieck, K. H. *Astrophys. J.* **209**, 653–662 (1976).

High-pressure (Mg, Fe)₂SiO₄ phases in the Tenham chondritic meteorite

THE high-pressure phase transformations of olivine (Mg, Fe)₂SiO₄ are of fundamental importance to the dynamics of the mantle^{1–3}. Experiments⁴ have demonstrated that at high pressure olivines transform to a solid solution series of (Mg, Fe) spinel (γ phase), while in olivines with Mg/Mg+Fe ratios greater than 0.85, 'modified' spinels with an orthorhombic structure (β phase) are formed⁵. The β phase in pure Mg₂SiO₄ is considered to have a stability field at pressures intermediate between those of olivine and the γ spinel phase, although the possibility that it forms as a metastable quench product from γ spinel has not been entirely ruled out. A central issue is the mechanism and hence the kinetics of these transformations but no such experimental work has been carried out. Details of the mechanisms involved may be inferred from experimental results, but available natural samples of high-pressure (Mg, Fe)₂SiO₄ polymorphs are confined to shock-produced

Table 1 Electron microprobe analysis of purple ringwoodite and olivine from the Tenham chondrite

	Ringwoodite (wt %)	Olivine (wt %)
Na ₂ O	0.00	0.00
MgO	39.24	39.73
Al ₂ O ₃	0.00	0.00
SiO ₂	38.66	38.67
K ₂ O	0.00	0.00
CaO	0.00	0.00
TiO ₂	0.00	0.00
Cr ₂ O ₃	0.00	0.00
MnO	0.24	0.47
FeO	22.75	22.83
NiO	0.00	0.00
Total	100.89	101.70
Formula	Mg _{1.51} Fe _{0.49} Si _{1.0} O ₄	Mg _{1.52} Fe _{0.49} Si _{1.0} O ₄

Analyses by S. O. Agrell.

phases in chondritic meteorites (for example, Tenham, Coorara, Catherwood and Coolamon^{6–10}). Binns *et al.*⁶ first reported the X-ray diffraction evidence for the existence of natural γ spinel in the Tenham chondrite where it occurs as small ($\approx 100 \mu\text{m}$), purple grains within shock-produced veins. The natural γ spinel is termed ringwoodite. Shock-wave experiments by Jeanloz and Ahrens¹¹ have failed to produce a spinel phase, and recently Jeanloz¹² has claimed, from X-ray powder diffraction, IR and optical data that ringwoodite is a complex aggregate of goethite, majorite garnet and γ iron, mis-identified as the spinel phase by previous workers. Here we present definitive evidence from electron-probe microanalysis, transmission electron microscopy and diffraction on the true nature of the purple grains (termed ringwoodite) in the Tenham chondrite.

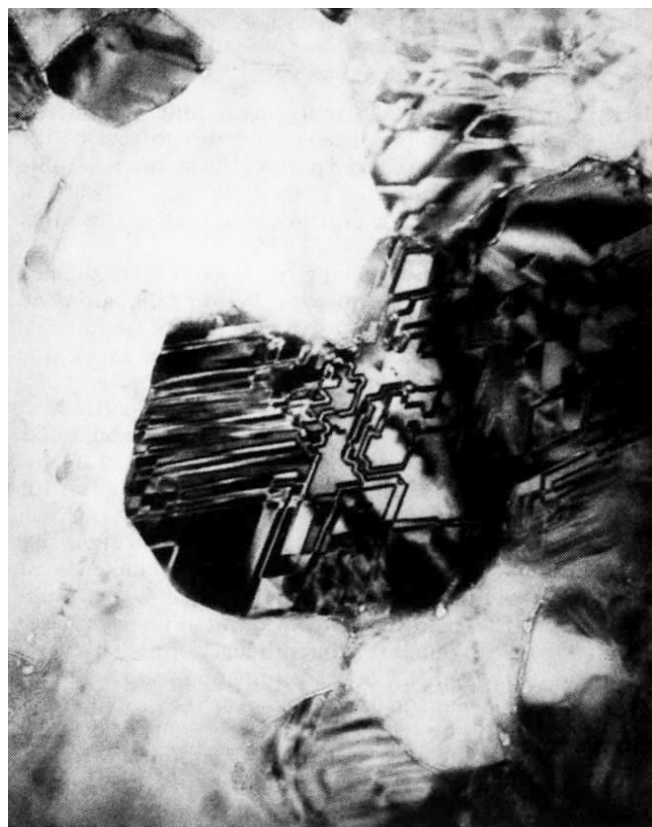


Fig. 1 Transmission electron micrograph of one crystal from a polycrystalline intergrowth formed by the transformation of olivine to γ spinel. The antiphase and twin boundaries within the crystal are due to the inversion of γ spinel to the β 'modified' spinel phase. Other crystals in this section are tilted out of contrast (magnification, $\times 140,800$).

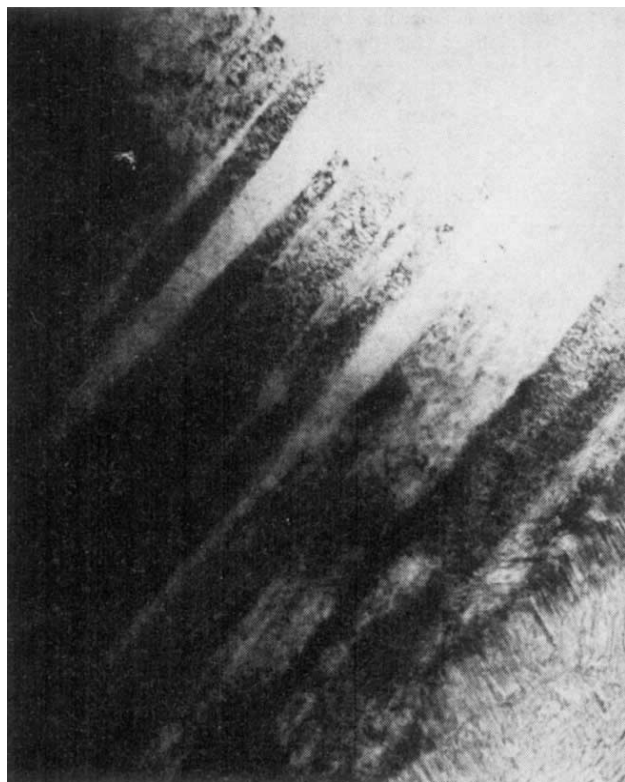


Fig. 2. Transmission electron micrograph showing the second type of morphology produced by the olivine to γ spinel transformation. The twin-related plates lie on $\{111\}$ planes of the γ spinel structure. Subsequent inversion to the β 'modified' spinel phase has produced the very fine antiphase domains within the plates (magnification $\times 60,200$).

Electron microprobe analysis of ringwoodite indicates a composition identical, within analytical error, to that of the olivine in the relatively unshocked parts of the meteorite (Table 1). This agrees with the results of Binns *et al.*⁶. Analyses do not vary significantly from grain to grain which in itself suggests that a complex aggregate is unlikely.

Optically thin sections containing ringwoodite were thinned for transmission electron microscopy by ion bombardment. Observations were recorded after various stages of the thinning process and a complete traverse made through the shock vein into the ringwoodite. Correlation of thinned areas with optical micrographs ensured that TEM observations were carried out on the purple grains where the composition has been confirmed. Electron petrographical observations of the microstructure and phase assemblages of the shock vein will be reported in full elsewhere.

Transmission electron microscopy and diffraction show that the purple ringwoodite consists of a polycrystalline spinel aggregate partially inverted to the β 'modified' spinel phase (Figs 1 and 2). The microstructures within the grains indicate that the β phase is an inversion product formed by rapid quenching of the γ spinel. The transformation from olivine to the original γ spinel has apparently taken place by two different mechanisms leading to spinel grains with distinctly different morphologies: first, as equant subregular intergrown crystals (grain size $\approx 1,000$ Å) (Fig. 1) and second as twin-related parallel plates on $\{111\}$ planes of the spinel structure (Fig. 2). The former morphology indicates an independent nucleation and growth mechanism while the latter suggests a cooperative shear mechanism with some of the characteristics of a martensitic transformation. As the transformation from olivine to spinel involves a change in the close packing of oxygen atoms, the mechanism may be dependent on the orientation of the original olivine crystal with respect to the shock-wave front. No

previous observations of possible mechanisms have been made, although Sung and Burns^{3,13} have emphasised nucleation and growth in attributing deep-focus earthquake genesis to this transformation. It is questionable whether the possibility of a martensitic mechanism would enable the necessary overpressures to be generated for an implosive transformation.

The subsequent inversion of γ spinel to the orthorhombic β 'modified' spinel phase on quenching requires no change in the oxygen packing and involves a redistribution of the cations^{5,14}. Thus the original γ spinel grain boundaries and the twin-related oxygen packing distribution (Fig. 2) are preserved. The relationship between the two structures introduces a very large number of translational and orientational variants leading to a complex anti-phase and twin domain structure within the grains (Figs 1 and 2). Electron diffraction patterns show splitting of the main spinel reflections due to the formation of twin domains of lower symmetry. Weak superlattice reflections are also observed which are consistent with the formation of the $a \times \sqrt{2}a \times a/\sqrt{2}$ body-centred orthorhombic cell of the β phase (where $a \approx 8.2$ Å, the lattice parameter of the γ spinel). The fine scale of the microstructure within the grains indicates that the nucleation of the β phase has occurred during a relatively rapid quench. There is no evidence in the microstructures or the diffraction patterns to suggest that the γ spinel phase was formed from olivine through any intermediate stage. The similarity between the β phase and γ spinel results in very similar X-ray powder patterns, that is predominantly that of spinel, with some extra lines. IR spectra will also indicate some similarities to spinel although with a significantly lower structural symmetry. Jeanloz' recent X-ray and IR data¹² are consistent with our observations that the purple grains are a γ spinel- β phase intergrowth.

The quench texture of the β phase again raises the question of its thermodynamic stability. Direct experimental evidence can be obtained by either studying the transformation *in situ* at high pressures, or more easily by observing the microstructures preserved in quenched experiments. Such TEM studies have yet to be carried out. Note, however, that the γ - β transformation has many of the characteristics of the formation of metastable transitional phases in non-equilibrium processes. A direct transformation from γ spinel back to olivine on pressure and temperature reduction involves a major structural change with a large activation energy. Metastable phases formed as kinetically more favourable alternatives are invariably modifications of the parent phase and involve minor structural reorganisation and a reduction in translational symmetry¹⁵. The β phase has these characteristics.

The purple grains in the Tenham chondrite described here represent the first reported natural occurrence of β -(Mg, Fe)₂SiO₄. Following Ringwood and Major⁴, we suggest that this β phase in nature be termed Wadsleyite.

We thank Drs S. O. Agrell and J. D. C. McConnell for discussions and I. C. Marshall for specimen preparation. Financial support was provided by the NERC.

A. PUTNIS
G. D. PRICE

Mineral Sciences Group,
Department of Mineralogy and Petrology,
Downing Place,
Cambridge, UK

Received 11 May; accepted 6 June 1979.

1. Ringwood, A. E. *Earth planet. Sci. Lett.* **14**, 233-241 (1972).
2. Turcotte, D. L. & Schubert, G. *J. geophys. Res.* **76**, 7980-87 (1971).
3. Sung, C. M. & Burns, R. G. *Tectonophysics* **31**, 1-32 (1976).
4. Ringwood, A. E. & Major, A. *Phys. Earth planet. Int.* **3**, 89-108 (1970).
5. Moore, P. B. & Smith, J. V. *Phys. Earth planet. Int.* **3**, 166-177 (1970).
6. Binns, R. A., Davis, R. J. & Reed, S. B. *J. Nature* **221**, 943-944 (1969).
7. Binns, R. A. *Phys. Earth planet. Int.* **3**, 156-160 (1970).
8. Smith, J. V. & Mason, B. *Science* **168**, 832-33 (1970).
9. Coleman, L. C. *Can. Miner.* **15**, 97-103 (1977).
10. Steele, I. M. & Smith, J. V. *Lun. planet. Sci.* **9**, 1101-1103 (1978).
11. Jeanloz, R. & Ahrens, T. J. *Geol. Soc. Am. Abstr. Prog.* **9**, 1036 (1977).
12. Jeanloz, R. *Lun. planet. Sci.* **10**, 619-621 (1979).
13. Sung, C. M. & Burns, R. G. *Earth planet. Sci. Lett.* **32**, 165-170 (1976).
14. Sung, C. M. & Burns, R. G. *Phys. Chem. Miner.* **2**, 177-197 (1978).
15. Putnis, A. *Science* **193**, 417-418 (1976); *Phys. chem. Miner.* **3**, 183-197 (1978).

New phyllosilicate types in a carbonaceous chondrite matrix

CARBONACEOUS chondrites provide valuable information as they are the least altered examples of early Solar System material¹. The matrix constitutes a major proportion of carbonaceous chondrites. Despite many past attempts, unambiguous identification of the minerals in the matrix has not been totally successful². This is mainly due to the extremely fine-grained nature of the matrix phases. Recently, progress in the characterisation of these phases has been made by electron diffraction studies^{3,4}. We present here the direct observation, by high resolution imaging, of phases in carbonaceous chondrite matrices. We used ion-thinned sections from the Murchison C2(M) meteorite for transmission electron microscopy. The Murchison matrix contains both ordered and disordered intergrowths of serpentine-like and brucite-like layers. Such mixed-layer structures are new types of layer silicates.

For an unknown mineral, electron diffraction patterns from a number of principal crystallographic zones are a sufficient (although not necessary) means of identification. However, selected area diffraction patterns of phyllosilicates from the Murchison matrix are of poor quality because of the small sizes of individual grains. Good images, on the other hand, can be obtained. Interpretation of these images is based on previous investigations of Murchison⁵, and on images from related high-resolution transmission electron microscopy (HRTEM) studies of terrestrial phyllosilicates⁶⁻⁹. Phyllosilicate structure types may be identified by their characteristic (001) basal spacings when viewed parallel to their sheets¹⁰. For example, serpentines show a basal spacing of 7.3 Å, while micas and talcs have a layer repeat of 10 Å (ref. 11).

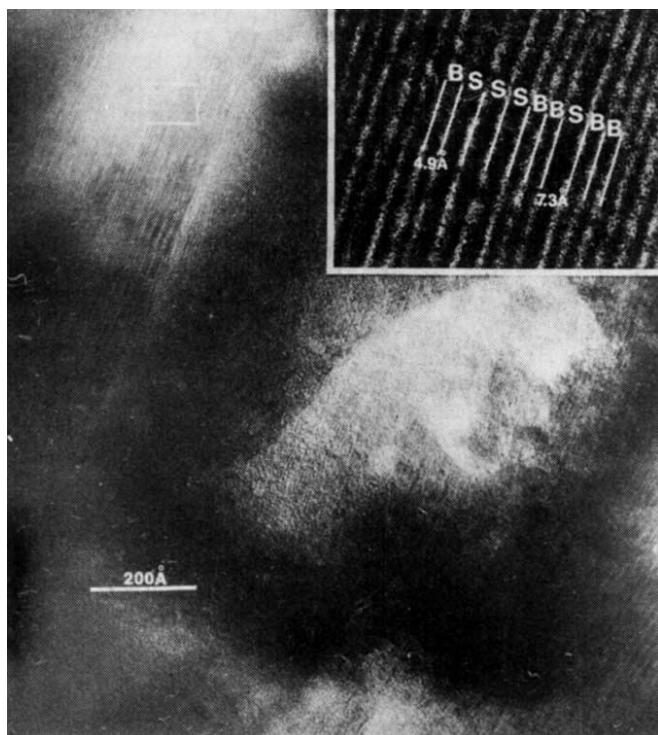


Fig. 1 A HRTEM image showing basal fringes in several mixed layer silicate crystals within the matrix of the Murchison carbonaceous chondrite. The inset is an enlargement of a portion of the same area which reveals a disordered arrangement of serpentine-(S) and brucite-type (B) layers.

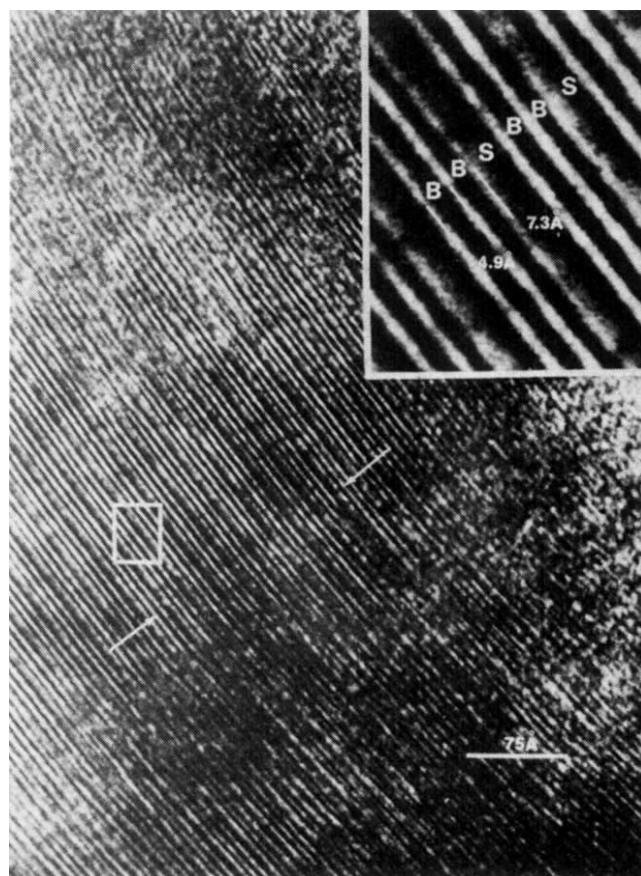


Fig. 2 A HRTEM image of layer silicate ordering in the Murchison matrix. An ordered sequence of 10 SBB units is indicated between the arrows. The inset is an enlargement of the sequence shown in the white rectangle.

The matrix material in Murchison consists of amorphous material¹⁷, and crystalline regions ranging in width from a few thousand to less than one hundred Ångströms. Discrete crystals occur surrounded by amorphous material; crystalline material also occurs as intergrowths in sub-parallel orientations, or as apparently aggregated, but randomly orientated groups of crystals. A considerable fraction (~50%) of the crystalline phases are phyllosilicates.

The extremely small grain size and character of the matrix crystals have also been confirmed by selected area microdiffraction using high-resolution scanning transmission electron microscopy (STEM). (Using a 20-μm limiting aperture in a STEM (Vacuum Generators HB5) equipped with a field-ion emission gun, the effective beam size is 10 Å. Selected area microdiffraction patterns of many phyllosilicate orientations were obtained with beam movement of <100 Å across the Murchison matrix.) Interpretation of microdiffraction patterns having periodicities of ~10 Å or greater is uncertain (J. M. Cowley, personal communication), but is sufficient to confirm the observations obtained by HRTEM imaging.

Figure 1 is a HRTEM photograph of part of the Murchison C2 carbonaceous chondrite showing typical phyllosilicate basal fringes in several crystals that are orientated in sub-parallel directions. The inset is a magnified portion of the same area, showing a disordered arrangement of 7.3 and 4.9 Å fringes representing layers of different types. (Fringe spacings are measured with a relative error of ~10%.) Fringe spacings of ~10 Å have also been observed in other portions of the Murchison matrix¹².

Basal fringe spacings of 7.3 Å are typical of minerals from the serpentine group. This group contains minerals ranging in composition from the pure Mg rich end member, serpentine, to the Fe, Mg, Al-silicates of the septeclorite series¹¹. Here we refer to the 7.3 Å fringes as representing serpentine layers in order to imply a structural type, rather than a compositional variety. The sheet hydroxides brucite, Mg(OH)₂, and gibbsite, Al(OH)₃, have basal spacings of 4.9 Å (ref. 11). However, the low total aluminium content of the Murchison matrix⁵ excludes gibbsite as a likely species in these layered intergrowths. Interpretation of the images as resulting from alternating layers of serpentine (S) and brucite (B) is consistent with compositional data⁵ and the morphology of phyllosilicates at high resolution⁶⁻¹⁰. Serpentine-brucite associations are also known to occur in terrestrial rocks¹³ and are consistent with the optical and X-ray studies of Fuchs *et al.*⁵.

As far as we know coherently intergrown serpentine- and brucite-type layers have not been described previously. Our observations of such intergrowths are a consequence of improvements in instrumental resolution and techniques as applied to meteoritic phyllosilicates. Assuming that our interpretation of the phyllosilicate phases is correct, it is possible to define several types of intergrowths. For example, we have observed (1) disordered sequences of S and B (Fig. 1), and (2) ordered sequences of ...SBBSBB... (Fig. 2). The origin and stability ranges, if any, of such new types of intergrowths are yet to be determined. Such information may prove to be genetically significant and potentially can provide information about the origins of carbonaceous chondrite matrices. Future experimental studies could well be directed towards establishing pressure-temperature stability ranges for such mixed layer assemblages.

A remarkable feature of the phases shown in Fig. 2 is the nature and extent of the phyllosilicate ordering. Repeats of serpentine and brucite occur continuously for more than 300 Å (...SBBSBB...). Following the procedures of Veblen and Buseck¹⁴, calculations for ordered SBB sequences in the Murchison matrix indicate that 20 repeating units have a probability of 1.57×10^{-15} of occurring randomly.

Experimental work on the serpentine system¹⁵ provides a limited, but, nevertheless, useful guide to the conditions of phyllosilicate formation in the Murchison matrix. Intimate mixtures of brucite-like and serpentine-like layers and their cylindrical morphology^{6,10} suggest that the 7.3 Å fringes arise from a Mg-rich serpentine such as chrysotile. An excess of Mg and relatively low silica activity at the time of matrix formation could possibly promote the growth of brucite layers. Alternatively, moderate silica activity and slightly lower temperatures might encourage the growth of brucite and iron-rich serpentines (septeclorites).

The data of Hemley *et al.*¹⁵ suggest that chrysotile and brucite can exist stably between 250 and 375 °C at 1 kbar water pressure. The effects of Fe and Al in the serpentine system are competitive: Fe reduces the temperature range slightly, while the upper stability limit for serpentine is raised by the presence of Al¹³. Specific compositional data are not yet available for the Murchison serpentine-brucite layers. However, indications of compositional variations from Mg-rich serpentine have been obtained in other C2 carbonaceous chondrites¹⁶ and suggest that Fe and Al occur in the layered intergrowths.

A more pervasive influence on the serpentine-brucite stability range may be pressure. The effect of pressure on serpentine stability is not well known. However, at the pressures assumed for the presolar nebula (10^{-2} – 10^{-3} atmospheres⁵), the chrysotile-brucite stability field would occur at considerably lower temperatures. Thus, 250–375 °C is a probable upper temperature limit for some of the phyllosilicates in the Murchison matrix.

We thank Drs David Veblen, Tom McKee and Gerry Gibbs for useful discussions. The microscopy was performed at the electron microscope facility in the Centre for Solid State Science at Arizona State University, where John Wheatley and Mary-Lynn Formanack were of assistance. Carleton Moore and the

Centre for Meteorite Studies provided the Murchison sample. Financial support was provided by NASA grant NSG 9068.

IAN D. R. MACKINNON
PETER R. BUSECK

Departments of Chemistry and Geology,
Arizona State University,
Tempe, Arizona 85281

Received 20 March; accepted 30 May 1979.

1. Research News, *Science*, **202**, 203–204 (1978).
2. Nagy, B. *Carbonaceous Meteorites* (Elsevier, New York, 1975).
3. Barber, D. J. *Meteoritics* **12**, 172–173 (1977).
4. Müller, W. F., Kurat, G. & Kracher, A. *Meteoritics* **12**, 322 (1977).
5. Fuchs, L. H., Olsen, E. & Jensen, K. J. *Smith. Contr. Earth Sci.*, No. 10 (1973).
6. Yada, K. *Acta Crystallogr.* **23**, 704–707 (1967).
7. Yada, K. & Iishi, K. *Am. Miner.* **62**, 958–965 (1977).
8. Iijima, S. & Buseck, P. R. *Trans. Am. geophys. Un.* **58**, 525, 1977.
9. Iijima, S. & Buseck, P. R. *Acta Crystallogr.* **A34**, 709–719 (1978).
10. Veblen, D. R. & Buseck, P. R. *Science* (in the press).
11. Deer, W. A., Howie, R. H. & Zussman, J. *Rock Forming Minerals* Vol. 3 (Longmans, London, 1962).
12. Mackinnon, I. D. R. & Buseck, P. R. *Proc. 10th Lunar Sci. Conf.* (in the press).
13. Wicks, F. J. & Whittaker, E. J. W. *Can. Miner.* **15**, 459–488 (1977).
14. Veblen, D. R. & Buseck, P. R. *Am. Miner.* (in the press).
15. Hemley, J. J., Montoya, J. W., Shaw, D. R. & Luce, R. W. *Am. J. Sci.* **277**, 322–352 (1977); 353–383 (1977).
16. McKee, T. R., Moore, C. B. & Chang, S. *Lunar Planet. Sci.* **10**, 807–809 (1979).
17. McKee, T. R. & Moore, C. B. *Proc. 10th Lunar Sci. Conf.* (in the press).

An experimental measurement of cation diffusion in almandine garnet

COMPOSITIONAL zoning in minerals is a disequilibrium phenomenon which is in principle potentially useful for interpreting the thermal history of the host rock. Almandine-rich garnets from low and medium grade regionally metamorphosed rocks are commonly zoned¹⁻⁸. The manganese content typically varies from core to rim, being high in the centre and low at the rim, producing a bell-shaped profile, whereas Fe is antipathetic to Mn. With the exception of the diffusion models of Anderson and Buckley⁹, such zoning has generally been explained in terms of fractionation-depletion processes²⁻⁴ where diffusion in the growing mineral is assumed to be negligible. However, in garnets from high grade rocks^{8,10,11}, zoning is usually weak, absent or reversed, and recent authors^{8,12,13} have favoured the homogenisation of pre-existing or developing zoning by volume diffusion. To evaluate the importance of cation diffusion to garnet zoning, Fe–Mn interdiffusion coefficients have been measured over a restricted temperature range.

Interdiffusion experiments were performed between almandine-rich garnet (100–400 µm in diameter) of approximate composition (Mg_{0.93}Fe_{2.07})Al₂Si₃O₁₂ and a sinter of Mn₃Al₂Si₃O₁₂. Pseudobinary couples comprising garnet and sinter (1:5 weight ratio) were packed in iron crucibles, and sealed in alumina tubes with alumina cement. One atmosphere diffusion runs were performed at temperatures of 822, 915, 934 and 1,002 °C for periods for 300–501 h.

Several garnets of approximately the same size were recovered from each experiment and mounted together for analysis. Only light grinding for a few seconds on 600 grade SiC paper was necessary to reach the centres of the crystals. Provided that the interdiffusion region does not account for more

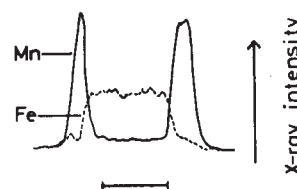


Fig. 1 Fe–Mn interdiffusion in almandine garnet (specimen 117: 915 °C, 501 h) variation of Fe and Mn X-ray intensity along traverse through the centre of a euhedral crystal. Scale bar, 80 µm.

than 20% of the garnet radius (r), the measured diffusion profile will not be significantly distorted if the exposed face is not more than $0.3r$ from the true centre. Samples were polished with diamond paste, coated with carbon, and examined by Cambridge Instruments Geoscan Microprobe.

The interdiffusion regions were analysed with a 'stepping interval' of 2 μm . Operating and data correction procedures are described in detail elsewhere¹⁴. As it was not possible to identify how close the exposed face was to the crystal centre, errors from misalignment were minimised by selecting only the largest crystals with the smallest interdiffusion regions (always less than $0.3r$). Interdiffusion coefficients ($\tilde{D}$) were calculated by the Boltzmann–Matano method¹⁵, and in common with other systems where only one member of the diffusion couple is available for analysis¹⁶, the Matano interface was assumed to coincide with the crystal rim. The maximum error in $\tilde{D}$ resulting from misalignment is thought to be $\pm 15\%$, and this has been incorporated in the final uncertainty of the data.

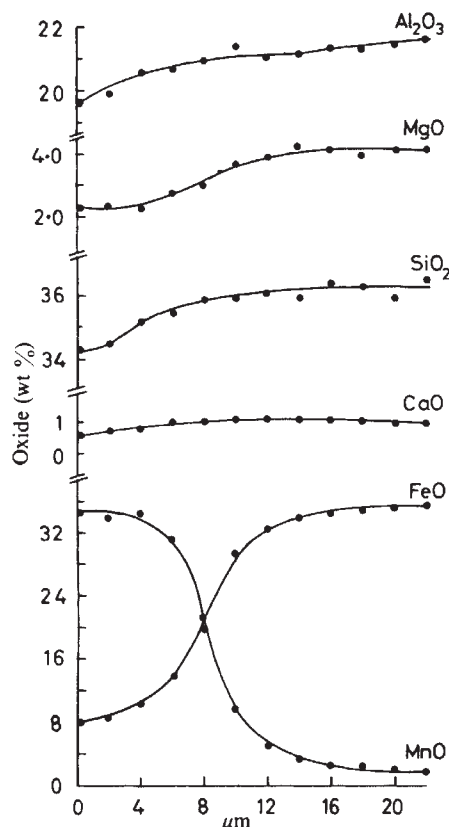


Fig. 2 Typical profiles of Al, Mg, Si, Ca, Fe and Mn in the interdiffusion region of garnet specimen 117 (915 °C, 501 h).

Although the experiments were performed outside the stability field of the garnet, there was no sign of breakdown, or subsequent surface growth. Measurable diffusion profiles were obtained at temperatures above 900 °C. The interdiffusing elements (Fe and Mn) exhibit complementary profiles (Fig. 1) and a detailed analysis is presented in Fig. 2. The shape of the Fe and Mn composition profiles warrants some comment. The flattening of the curves near the interface is probably due to the use of a sinter^{16,17} rather than single crystal spessartine. The intersection of the Fe and Mn profiles some 8 μm inside the crystal is of no significance as the cross-over point should not necessarily coincide with either the Matano interface or the original interface^{18,19}. It is possible that some interdiffusion of Mn–Mg and Mn–Ca has occurred but the effect on the location of the Matano interface is considered negligible. To a good approximation, the description of the system as an Fe–Mn pseudobinary is valid.

Results for $\tilde{D}$ (Fe–Mn), based on a minimum of eight crystals, are shown in Fig. 3. The composition-dependence of the interdiffusion coefficients may be described by

$$\tilde{D} = D_{\text{Mn}}^* \exp(\beta X_{\text{Mn}}) \quad (1)$$

where D_{Mn}^* is the tracer diffusion coefficient of Mn in Mn-free garnet, X_{Mn} the content of Mn in wt% (to a maximum of 33.1 wt% for the complete replacement of Fe by Mn) and β the composition factor. Over the temperature range 915–1,002 °C, D_{Mn}^* increased from $3.7 \times 10^{-18} \text{ m}^2 \text{ s}^{-1}$ to $6.6 \times 10^{-18} \text{ m}^2 \text{ s}^{-1}$, whilst $\beta = 0.09 \pm 0.01$.

The temperature-dependence of the interdiffusion process may be obtained from an Arrhenius plot of the form

$$\tilde{D}_{\text{Mn}} = D_0 \exp(-Q/RT) \quad (2)$$

where $\tilde{D}_{\text{Mn}}$ is the interdiffusion coefficient at constant composition, D_0 the pre-exponential factor, and Q the activation energy (in kJ mol^{-1}). Results at 10 and 15 wt% Mn may be described by $\tilde{D}_{10} = 1.07 \times 10^{-13} \exp[-94(\pm 16) \text{ kJ}/RT] \text{ m}^2 \text{ s}^{-1}$, and $\tilde{D}_{15} = 2.01 \times 10^{-14} \exp[-72(\pm 25) \text{ kJ}/RT] \text{ m}^2 \text{ s}^{-1}$ respectively.

There is good agreement between the present $\tilde{D}$ (Fe–Mn) results and data from a recent petrological study. Hensen²⁰ examined the reaction of garnet (of composition $\text{Mg}/(\text{Mg} + \text{Fe}^{2+}) \approx 0.28$) and cordierite at 1,000 °C and 10 kbar, and found that some garnet grains did not homogenise in 305 h. With grains of 10–40 μm diameter, diffusion distances of 5–20 μm would be required to produce uniform compositions. Assuming a diffusion distance, x , of $15(\pm 5) \mu\text{m}$, in a time, t , of 305 h, then in this case $\tilde{D}$ (Fe–Mg) may be estimated from an approximate solution of the diffusion equations¹⁵

$$x^2 \approx KDt \quad (3)$$

where $K \approx 4$ and is a constant²¹. The calculated interdiffusion coefficient of $5(\pm 3) \times 10^{-17} \text{ m}^2 \text{ s}^{-1}$ compares well with the value of $2.3 \times 10^{-17} \text{ m}^2 \text{ s}^{-1}$ obtained here for $\tilde{D}$ (Fe–Mn) at 15 wt% Mn. By defining $\tilde{D}[\text{Fe}-(\text{Mn}, \text{Mg})]$ as the interdiffusion coefficient of Fe with either Mn or Mg, an estimate for this parameter may be obtained by combining the present data for $\tilde{D}$ (Fe–Mn) and the result deduced from Hensen for $\tilde{D}$ (Fe–Mg). At compositions of ~ 13 wt% (Mn, Mg) in silicate garnets $\tilde{D}[\text{Fe}-(\text{Mn}, \text{Mg})] = 6.77 \times 10^{-12} \exp[-132(\pm 45) \text{ kJ}/RT] \text{ m}^2 \text{ s}^{-1}$. Unfortunately, experimental data are available for only a restricted temperature range, and it is dangerous to extrapolate coefficients to temperatures outside the range of measurement since the diffusion regime and Q may change²². A direct assessment of diffusion rates at significantly lower temperatures must await more sophisticated experiments.

Evidence for volume diffusion in natural garnets has recently been cited^{8,12,13}, but in general, the interpretation of zoning is not easy. In a survey of garnets from west central Massachusetts, Tracy, Robinson and Thompson⁷ found they were unable to distinguish fully between garnet zoning due to continuous prograde, or retrograde reactions, or exchange with Fe–Mn–Mg inclusions, or diffusion within the garnet. Other authors^{8,12} have suggested that all garnets develop zoning by a fractionation-depletion process during growth, which is subsequently homogenised by diffusion, particularly in the highest grades. From

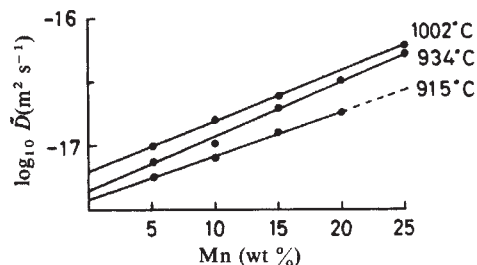


Fig. 3 Composition-dependence of $\tilde{D}$ (Fe–Mn) in almandine garnet at one atmosphere for the temperature range 915–1,002 °C.

estimates of apparent diffusion distances (x), and temperatures of the metamorphic zones, Yardley⁸ calculated Q for diffusion in garnet to be 293–418 kJ mol⁻¹. In principle this is a reasonable approach but x may be hard to assess. In the highest grades, homogenised garnets may unequivocally be attributed to volume diffusion if the earlier zoning is recorded, for example, in the composition of inclusions¹². However, in the lower grade rocks effective diffusion distances may be ambiguous since garnets from a single sample often exhibit a range of zoning profiles^{13,23}. It is possible, therefore, that Yardley underestimated x for his samples 146 and 149, and by using plausible alternatives, activation energies as low as 154 kJ mol⁻¹ are obtained.

The present experiments confirm that diffusion does occur in garnets, that $\bar{D}[\text{Fe}-(\text{Mn}, \text{Mg})]$ is typically 10⁻¹⁷ m² s⁻¹ at temperatures around 950 °C, and activation energies are probably in excess of 100 kJ mol⁻¹. In environments of geological interest interdiffusion rates will be sensitive to (T , P), composition gradients, bulk composition and possibly shearing stress¹², or the presence of a fluid, and thus the detailed thermal history may be difficult to elucidate.

I thank the SRC for financial support and Drs S. C. Elphick, C. M. Graham and R. G. J. Strens for helpful comments.

R. FREER*

School of Physics,
University of Newcastle upon Tyne,
Newcastle, UK

Received 4 April; accepted 11 June 1979.

* Present address: Grant Institute of Geology, University of Edinburgh, UK.

1. Harte, B. & Henley, K. J. *Nature* **210**, 689–692 (1966).
2. Hollister, L. S. *Science* **154**, 1647–1651 (1966).
3. Atherton, M. P. & Edmunds, W. M. *Earth planet. Sci. Lett.* **1**, 185–193 (1966).
4. Atherton, M. P. *Contr. Min. Petrol.* **18**, 347–371 (1968).
5. Hollister, L. S. *Bull. Soc. Am.* **80**, 2465–2494 (1969).
6. Cooper, A. F. *J. Petrol.* **13**, 457–492 (1972).
7. Tracy, R. J., Robinson, P. & Thompson, A. B. *Am. Miner.* **61**, 762–775 (1976).
8. Yardley, B. W. D. *Am. Miner.* **62**, 793–800 (1977).
9. Anderson, D. E. & Buckley, G. R. *Contr. Min. Petrol.* **40**, 87–104 (1973).
10. Blackburn, W. H. *Can. Miner.* **9**, 691–698 (1969).
11. Grant, J. A. & Weiblen, P. W. *Am. J. Sci.* **270**, 281–296 (1971).
12. Woodsworth, G. J. *Can. Miner.* **15**, 230–242 (1977).
13. Anderson, D. E. & Olimpio, J. C. *Can. Miner.* **15**, 205–216 (1977).
14. Freer, R. thesis, Univ. Newcastle upon Tyne (1978).
15. Crank, J. *The Mathematics of Diffusion*, 2nd ed (Oxford University Press, 1975).
16. Buening, D. K. & Buseck, P. R. *J. geophys. Res.* **78**, 6852–6862 (1973).
17. Rigby, E. B. & Cutler, I. B. *J. Am. ceram. Soc.* **48**, 95–99 (1965).
18. Yurek, G. J. & Schmalzried, H. *Ber. Bunsenges. phys. Chem.* **78**, 1379–1386 (1974).
19. Appel, M. & Pask, J. A. *J. Am. ceram. Soc.* **54**, 152–158 (1971).
20. Hensen, B. J. A. *Rep. Dir. Geophys. Lab.*, Yb. **71**, 418–421 (1971).
21. Zaplatynsky, I. *J. Am. ceram. Soc.* **45**, 28–31 (1962).
22. Wagner, J. B. *Atomic Diffusion in Semiconductors* (ed. Shaw, D.) 543 (Plenum, London, 1973).
23. Kretz, R. *Can. Miner.* **12**, 1–20 (1973).

Palaeo-stress estimates in the Moine Thrust Zone, Eriboll, Scotland

STRESSES in the Earth's crust and mantle are of great interest to geologists and geophysicists, particularly the stresses associated with the formation of major tectonic structures. The measurement of quartz microstructural features reported here indicate the stresses during mylonitisation in the Moine Thrust Zone.

Recent studies^{1–5} have demonstrated that deformed and syntectonically recrystallised minerals in tectonites may be used to estimate the palaeo-stresses operative during their last deformation. The technique arises from the observation that in large-strain steady-state dislocation creep or hotworking of materials, a stable microstructure develops at low to moderate strains and subsequently remains approximately constant. Similar observations have been recorded for naturally deformed quartzites⁷.

The elements of the microstructure that are stress dependent are the dislocation density ρ , sub-grain size d and the size of the

recrystallised grains D . The following dependencies apply for metals⁶, ceramics⁸ and minerals^{9–11}:

$$\sigma = k\mu b\rho^{1/2} \quad \sigma = \frac{l\mu b}{d} \quad \sigma = md^{-0.7}$$

where σ is the differential stress, k , l and m are constants, μ the temperature compensated shear modulus and b the magnitude of the Burgers vector. In principle it is possible to make quantitative stress estimates if k , l and m have been determined experimentally for the mineral in question and if care is taken¹². The approximate values of l and m for quartz can be obtained from the observation¹³ that at 900 °C and a stress of ~100 MPa (1 kbar) the sub-grains in a deformed quartzite were approximately 5 μ m in diameter and the size of recrystallised grains was 20 μ m. Goetze⁹ has proposed a value of 3 for k . However, there is the possibility that ρ , d and D may have slight temperature dependencies¹², especially D in the case of minerals³. This must be resolved before accurate estimates can be made because there are normally large differences between the temperatures of experimental runs and those of crustal deformation. Temperature effects will lead to an overestimation of stresses and the values quoted below probably represent upper limits.

I have measured the variation in ρ , d and D in two mylonite zones in quartz-rich rocks in the Moine Thrust Zone at Eriboll, in North-West Scotland. The first was a narrow, 10-cm thick mylonite band in the Pipe Rock in the imbricate zone at Ben Heilam. A decreasing strain gradient exists away from the mylonite band with only slightly deformed Pipe Rock at ~25 cm from the mylonite band. The second zone was associated with the upper thrust plane (the Moine Thrust) of the Moine Thrust Zone. This outcrops in Alt Odhrsgaraidh. These mylonites gradually grade into the overlying Moine schists (see Soper and Wilkinson¹⁴). A sequence of specimens was collected from both localities. The first appearance of recrystallised grains around the old quartz grain margins was taken as the outer limit of the deformation producing the mylonites. Some low-strain deformation must exist beyond this limit to have produced the recrystallisation. However, it is probable that sub-grain sizes may not have equilibrated to the applied stress at these low strains and they were not measured.

The grain size was determined from optical micrographs of petrological thin sections. Dislocation density and sub-grain sizes were measured (diameters in the case of equidimensional sub-grains or widths if elongated) from electron micrographs taken from foils prepared from the above thin sections. All electron microscopy was performed in an AE1 EM7 high voltage transmission electron microscope operating at 1,000 kV.

The sub-grain size and the grain size, the mean of at least 100 measurements, in both sequences of specimens decreased into the mylonites and then remained constant. In the Pipe Rock, d decreased from 13.3 to 3.7 μ m (see Fig. 1) and D from 24.6 to 14.6 μ m. The corresponding decreases in the Moine mylonite zone were; d from 13.6 to 2.6 μ m and D from 42.9 to 18.6 μ m. Wherever possible, only sub-grains away from grain boundaries were measured because of a likely stress intensification in their vicinity. It was observed, in both mylonites, that the presence of mica led to a dramatic reduction in grain size and consequently grain size analyses were limited to mica-free areas. Unfortunately electron microscopy revealed that such areas, in optical micrographs, often contained small mica-grains. The dislocation densities varied erratically in both groups of samples from 2×10^7 to 1×10^8 intersections cm⁻² (averaging 5×10^7). This observation supports previous findings that care must be taken when using ρ for palaeostress estimates because of the ease with which it can be reset by small strains after the main deformation. The sub-grains have a greater stability as they are low energy structures containing geometrically necessary dislocations and were the most reliable palaeo-stress indicator in the specimens studied. The presence of mica adversely affected grain size estimates.

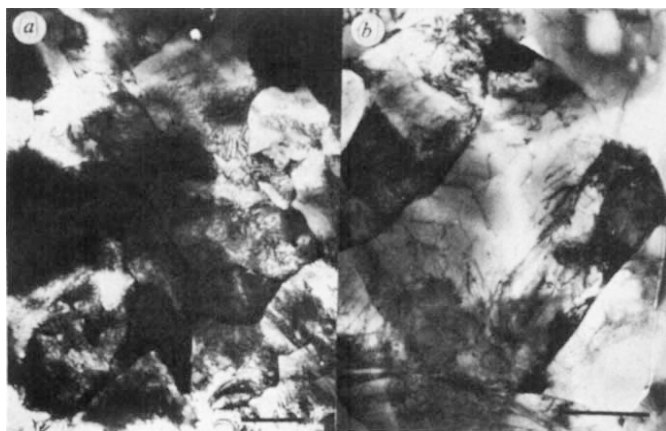


Fig. 1 Electron micrographs showing the decrease in sub-grain size in deformed quartz grains: *a* outside (scale bar, 10 µm); *b* inside (scale bar, 2 µm) a shear zone in the Pipe Rock.

The sub-grain sizes in both areas indicate a stress intensification by a factor between 3 and 4, into the mylonite zones. They indicate a similar low stress, 37 MPa (370 bar) at the edge of each deformation zone. The stresses increase to 128 MPa (1.28 kbar) in the Pipe Rock mylonites and to 180 MPa (1.8 kbar) in the Moine mylonites. The grain sizes indicate a smaller variation, for example, from 50 MPa (500 bar) to 100 MPa (1 kbar) for the Moine mylonites, but as stated previously little reliance can be placed on these values. The average dislocation density indicates a stress of ~35 MPa (350 bar) which is similar to the low stresses indicated by the sub-grains.

Shear zones and thrusts represent zones of inhomogeneous deformation and can be likened to shear band structures in metals, ceramics, polymers and soils^{15,16}. These bands mark zones of strain softening which may arise from mechanical or thermal effects. It has been suggested that mylonites also indicate zones of strain softening¹⁷. In the case of those studied there is no evidence for any marked shear heating and it is concluded that a mechanical softening mechanism was operating. I would expect the mylonite zones to initiate from local stress notches in the same way as shear bands initiate. Notches were not identified but would be expected as the Pipe Rock is inhomogeneous and in the case of the Moine Thrust there are marked changes in rock type. Therefore the propagation of the mylonite zone can be regarded as being similar to that of a shear band or a planar discontinuity in a cohesive material¹⁶. In such a model, which would apply to narrow zones, the stress gradient recorded by the sub-grains would be that at the tip of the propagating discontinuity. It would cause deformation in the surrounding material with that above and below frozen in after propagation. The softening due to the mylonite zone which develops leads to a relaxation of the localised stress concentration and also concentrates the deformation into the zone. Growth of the zone will occur if the flow rate in the mylonite cannot accommodate the overall imposed regional rate. The stresses required to produce the mylonite during the initiation and during the growth of a shear zone or ductile fault should be similar for a given rock type and deformation temperature. After the development of the required width of mylonite to accommodate the regional strain rate the adjacent country rock should play little further part in the deformation and the above gradients are preserved. Therefore the high stress estimates are likely to be those associated with the mylonite development and the low estimates should approximate to the regional background stresses to which the former relax. I suggest that the difference between the two represents the overall stress variation associated with the formation of the Moine Thrust Zone in the area studied. Note that the dislocation densities indicate a stress close to the relaxed stress. This may indicate

that, in this area, the dislocation densities have been reset to the regional stress after the mylonite formation.

In conclusion, I suggest that the quartz mylonites in the Moine Thrust Zone, at Eriboll, were produced by a localised stress of ~150 MPa (shear stress of 75 MPa or 750 bar) with the regional stress being ~35 MPa (shear stress of 17.5 MPa or 175 bar) which is the likely stress during thrusting after the mylonites formed.

I thank Dr R. Sibson and N. Shaw for helpful discussions, NERC and the USGS, Earthquake Hazards Reduction Program, contract 14-08-001-G-377 for funds.

S. WHITE

*Geology Department,
Imperial College,
London SW7, UK*

Received 27 March; accepted 11 June 1979

1. Kirby, S. H. & Raleigh, G. B. *Tectonophysics* **19**, 165 (1971).
2. White, S. Q. *Jl. geol. Soc. Lond.* **131**, 577 (1975).
3. Mercier, J. C., Anderson, D. A. & Carter, N. L. *Pageophysics* **115**, 119 (1977).
4. Twiss, R. J. *Pageophysics* **115**, 227 (1977).
5. Burg, J. P. & Laurent, P. *Tectonophysics* **47**, 15 (1978).
6. Takeuchi, S. & Argon, A. S. *J. mater. Sci.* **11**, 1542 (1976).
7. White, S. *Neues Jb. Min. Abh.* **123**, 237 (1975).
8. Evans, A. G. & Langdon, T. E. *Prog. mater. Sci.* **21**, 171 (1976).
9. Goetze, C. *Eos Trans. Am. geophys. Un.* **56**, 6 (1975).
10. Kohlstedt, D. L., Goetze, C. & Durham, W. B. in *The Physics and Chemistry of Minerals and Rocks* (ed. Strens, R. G. J.) **34** (Wiley, London, 1976).
11. Goetze, G. & Kohlstedt, D. L. *Contr. Miner. Petrol.* **59**, 296 (1977).
12. White, S. *Bull. Miner.* **102**, 210 (1979).
13. Ardell, A. J., Christie, J. M. & Tullis, J. A. *Cryst. Latt. Def.* **4**, 275 (1973).
14. Soper, N. J. & Wilkinson, P. *Scott. J. Geol.* **11**, 339 (1975).
15. Argon, A. S. *The Inhomogeneity of Plastic Deformation* 161 (ASM, Metals Park, Cleveland, 1973).
16. Palmer, A. C. & Rice, J. R. *Proc. R. Soc. A* **322**, 527 (1973).
17. White, S. *Phil. Trans. R. Soc.* **283**, 69 (1976).

Recent thrusting in the Appalachians

NUMEROUS orogenic episodes from late Precambrian to early Permian have shaped the Appalachian Mountains. Its last orogeny, the Allegheny orogeny, began before the end of the Lower Carboniferous and ended after the early Permian¹. Most of the Palaeozoic mountain belts of Europe and other continents were worn down long ago. It is still being discussed whether the Appalachians were constantly uplifted in post-Palaeozoic times or whether there was a Cenozoic rejuvenation of uplift²⁻⁴. The composition of the Mesozoic and Cenozoic Atlantic coastal sediments and their gentle dip away from the Appalachians imply that the mountains stood as a high broad arch despite two hundred millions of years of erosion¹. Geological evidence has been found and is reported here of thrusting during the past 12 yr at many locations in the Appalachian Plateau and Valley and Ridge province. Five sites have been selected in Tennessee, West Virginia, and Pennsylvania to show considerable horizontal shortening and associated vertical uplift that actually holds up the mountain ranges. The recent faulting and folding have been quantitatively studied and absolutely dated by offsets of boreholes that have been drilled for blasting operations in road-cuts.

Along the interstate highway 40 between Harriman and Rockwood, Tennessee folded and thrustured Pennsylvanian shales and sandstones are exposed striking N 45° E and dipping 10° towards south-east or north-west (Fig. 1). In this position the highway trends west-north-west and cuts across a syncline more than 15 km wide in the transition zone between the Valley and Ridge province and the Cumberland plateau following to the west.

West (10.5 km) of exit 338 (Westel road) on interstate highway 40 east, a thrust fault of supposedly Alleghenian age has recently been reactivated by an offset of 25 cm horizontal throw and 16 cm vertical movement (Fig. 2). The thrust plane strikes and dips N 42° E/32° SE and the hanging wall block is overthrust in the direction of N 80° W, being parallel to the

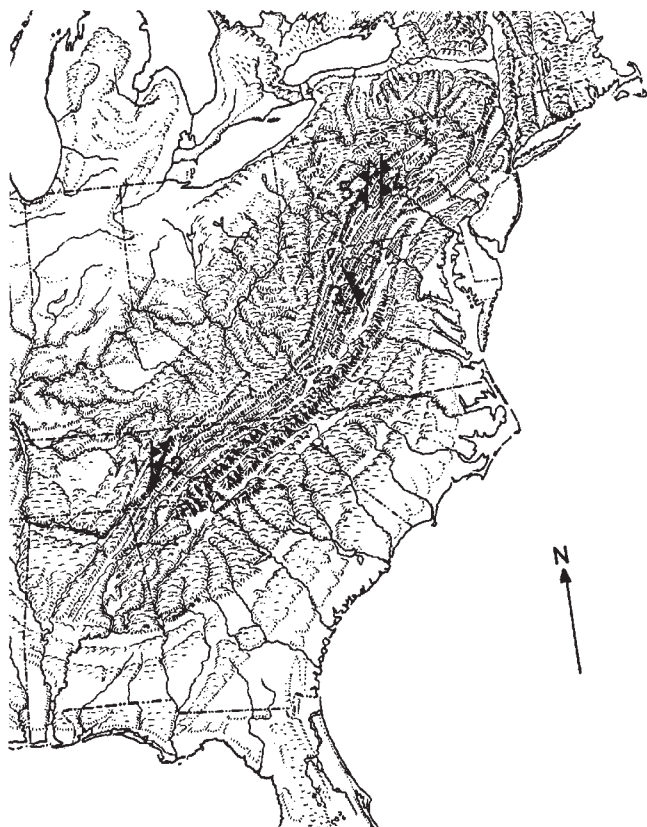


Fig. 1 Sites of recent thrusting in the Appalachian Plateau and Valley and Ridge province. Barbs on hanging wall blocks. Eastern part of US landform map by E. Raisz. Numbers of sites refer to Table 1.

orientation of the road-cut walls as can be seen by the displacements of the boreholes. Likewise, the direction of maximum horizontal stress that caused the thrusting seems to be $N 80^{\circ} W$ -orientated as no strain-slip of the rocks towards the open road-cut during the past 12 yr could be observed. The boreholes were drilled in summer 1966. The same thrust throw was manifested in April 1974 as in August 1978. Thus, for the period 1966–74 an annual average thrust-slip rate of 3.8 cm can be calculated. From 1974 to 1978 no distinct offset could be observed.

Four kilometres east of exit 338 on highway 40 east, boreholes were drilled during the summer of 1968. Here, on the eastern flank of the aforementioned syncline, Pennsylvanian sandstones and shales strike and dip $N 70^{\circ} E/20^{\circ} NW$ (Fig. 3). Highway 40 now is $N 40^{\circ} W$ -orientated and the displacements of the drill-holes at both sides of the road indicate thrust faulting along bedding planes, particularly at the transition from sandstones to shales. The direction of thrusting, however, is no longer strictly parallel to the road-cut walls. The boreholes are obliquely offset to the road on both sides of the highway revealing an east to $E 40^{\circ} S$ thrust direction of the hanging wall layers. The maximum horizontal throw recorded is 20 cm and the corresponding vertical movement is 7.5 cm. Some drill-holes are offset twice along different thrust (bedding) planes.

Considering the two exposures west and east of exit 338 on interstate highway 40 it is concluded that a recent stress field with a maximum compressive stress directed west-north-west has horizontally shortened an area of 14.5 km length by at least 45 cm and has induced vertical uplift of 7.5 cm to 16 cm during the past 12 or 10 yr, respectively (Fig. 4).

West (16 km) of Franklin, West Virginia and 3 km south-west of Oak Flat, West Virginia dark siltstones and shales of presumably middle Devonian age are exposed along highway 33 (Fig. 1). The sediments strike and dip $N 50^{\circ} E/15^{\circ} SE$. Boreholes were drilled there for blasting operations between June

and August 1973. Five years later the total horizontal displacements of the drill-holes along two bedding planes was about 50 cm. The hanging wall sediments were shifted to $W 15^{\circ} S$. As the direction of overthrusting almost corresponds with the strike of sediments no vertical uplift occurred. This considerable offset during such a short time-span seems to require other means for displacement in addition to tectonic stress. But the direction of thrusting is again parallel to the road-cut, although the inclination of the sediments towards the highway could have induced gravitational sliding at the north-western road-cut wall. A west-south-west-directed maximum horizontal stress has been suggested as the cause of the recent displacement of the boreholes whose major throw may have occurred shortly after the blasting. A major strain release by thrust-slip movements (soon after the blasting), is also indicated by the observations at site 1 (Fig. 1, Table 1).

Exposures of Devonian quartzites have been studied on highway 22, 14.5 km north-west of Millerstown, Pennsylvania. The layers are almost horizontal ($N 8^{\circ} E/5^{\circ} SE$) being close to the centre of a syncline within the Pennsylvanian Valley and Ridge province. According to borehole offsets only a minor thrusting in the direction of $N 80^{\circ} W$ has been observed. The horizontal throw amounts to 1.5 cm and vertical movements were calculated to be 0.1 cm. I do not know when the drilling and blasting for the construction of the road-cut occurred, but according to the state of the highway and the freshness of the exposed rocks along the road sides it seems that blasting was conducted no more than 10 yr ago.

The highway 22 continues northwestwards and intersects narrow folds of the eastern Allegheny plateau. About 6 km north-west of Port Matilda, Pennsylvania the highway 22/322 cuts across Mississippian and Pennsylvanian red beds of sandstones, siltstones and shales that strike and dip $N 60^{\circ} E/20^{\circ} NW$. Many displacements of drill-holes indicate overthrusting along bedding planes in the direction of $E 50^{\circ} S$. There are numerous displacements of minor throw within the sandstone sequences and there are a few thrusts of major throw along the sandstone/shale contact planes, where friction is reduced. Thus, the maximum thrusting with a horizontal throw of 14 cm and a vertical uplift of 5 cm has been measured of a sandstone over-riding a shale. Another thrust-fault was recorded with a horizontal throw of 5 cm and a vertical uplift of 1.8 cm. Along both sides of the highway 322, from 6 to 8 km north-west of Port Matilda, Pennsylvania, most of the 20° -NW-dipping thrust-faults indicate a movement of the hanging wall blocks in the direction of $E 50^{\circ} S$. But some borehole offsets at the south-west

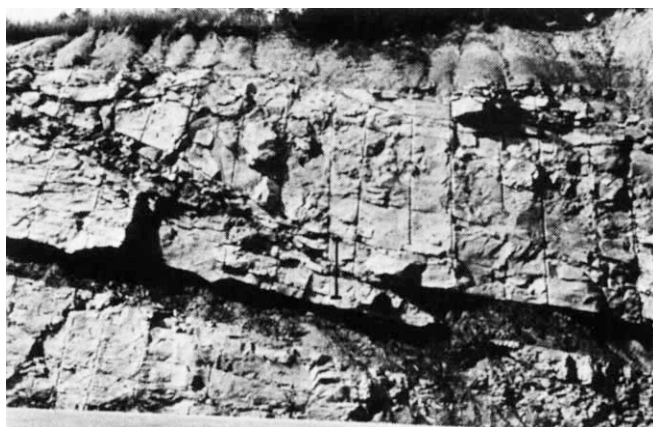


Fig. 2 Thrust-fault in Pennsylvanian sandstones intercalated by a dark shale layer, located between Rockwood and Harriman, Tennessee, 10.5 km west of exit 338 on interstate highway 40 east. Borehole offsets occurred after blasting in 1966. Displacements of 25 cm horizontal throw and 16 cm vertical uplift have been recorded in 1974. Direction of thrust motion is $N 80^{\circ} W$ and corresponds to the orientation of maximum horizontal stress. Scale bar, 1.83 m (6 ft).

Table 1 Sites of recent thrusting in the Appalachian Plateau and Valley and Ridge province

Location	Direction of thrusting and inferred direction of σH_{max}	Inclination of thrust plane	Horizontal throw	Vertical throw	Period of recent thrusting
1*. Rockwood-Harriman, Tennessee Interstate 40 east, 10.5 km west exit 338	N 80° W	32° SE	25 cm	16 cm	1966-74
2*. Rockwood-Harriman, Tennessee interstate 40 east, 4 km east exit 338	East to E 40° S	20° NW	20 cm	7.5 cm	1968-78
3*. 3 km south-west Oak Flat, West Virginia, highway 33	W 15° S	0°	50 cm		1973-78
4*. 14.5 km north-west Millerstown, Pennsylvania, highway 22	N 80° W	5° SE	1.5 cm	0.1 cm	about 10 yr
5*. 6-8 km north-west Port Matilda, Pennsylvania highway 322	E 50° S N 70° E	20° NW	14 cm 5 cm	5 cm 1.8 cm	about 10 yr

* Numbers correspond to sites in Fig. 1.

road side wall show strain-slip movements in direction of N 70° E towards the open road-cut. A recent maximum horizontal compressive stress which caused the thrusting is suggested to be north-west directed rather than west-north-west, as can be deduced from the strain-slip movements.

The drill-hole offsets were considered to have been caused mainly by blasting, as at site 1 where measurements have been repeated no displacement was observed during the past four years. However, there is evidence of updip thrusting after the blast operations and not of gravity controlled movements. The thrusting at sites 1 and 2 occurred in exactly the opposite directions due to the corresponding orientation of maximum horizontal stress (σH_{max}) at both locations. The change in the orientation of the road-cut at site 2 (N 40° W) in respect to the direction of the road excavation at site 1 (N 80° W) is not reflected by the orientation of the borehole displacements.

The energy released by the explosions is probably insufficient to displace rock masses of more than 50,000 m³ (site 5) updip on planes as one block.

Borehole offsets have been observed in southeastern Connecticut after blasting operations in 1970⁵. Measurements at one location have been repeated between 1970 and 1978 revealing that the initial displacement of 23 mm was followed by a period of creep at an average rate of 2.8 mm yr⁻¹. It is suggested that this permanent creep results from stresses acting on the Appalachians at present, rather than from palaeo-stresses that would have created one episodic rock movement during

or shortly after the blasting. As for site 1 it is suggested that thrusting will continue if σH_{max} reaches the critical value that was achieved some time between the blasting of 1966 and 1974.

If the borehole offsets on highway 322 between Port Matilda and Philipsburg, Pennsylvania and those on highway 22 between Lewistown and Millerstown, Pennsylvania are about 10 years old, we may apply the calculated amounts of annual uplift from the sites where recent thrusting has been studied to the Pennsylvanian Valley and Ridge province in general. Vertical movements of at least 2 cm yr⁻¹ can be inferred from the structural investigations on the highway 322 north-west of Port Matilda, Pennsylvania and of 0.1 cm on the highway 22 west-north-west of Millerstown, Pennsylvania. However, it has been

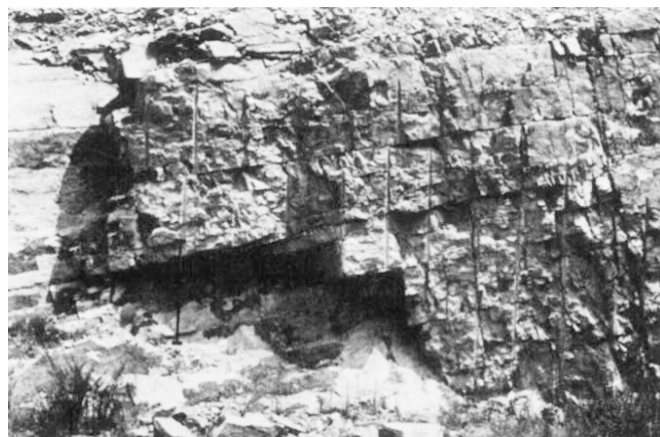


Fig. 3 Recent thrusting after drilling and blasting of boreholes in 1968 occurred along a bedding plane of Pennsylvanian sandstones, 4 km east of exit 338 on interstate highway 40 in eastern Tennessee. Scale bar, 1.83 m (6 ft).

frequently observed that recent thrusting is enhanced in 20°-30° inclined sediments and that thrusting decreases in more gentle or steeper dipping layers.

The crests of the main ridges of the Valley and Ridge province north-west of Harrisburg, Pennsylvania show small remnants of an old peneplain surface formed by an ancient Susquehanna river-system when the region was near base level. The Susquehanna river is now superimposed on the rejuvenated Valley and Ridge province that was regionally uplifted by about 300 m in relation to the present valley floor. If a continuous uplift of 2 cm yr⁻¹ is inferred for the recent and past vertical movements in that area, the present elevation of the ridge crests above the ancient peneplain will be achieved during a period of 15,000 yr only. However, at this point of investigations of recent thrusting and folding in the Appalachians we have to consider the thrust-slip rates as local phenomena.

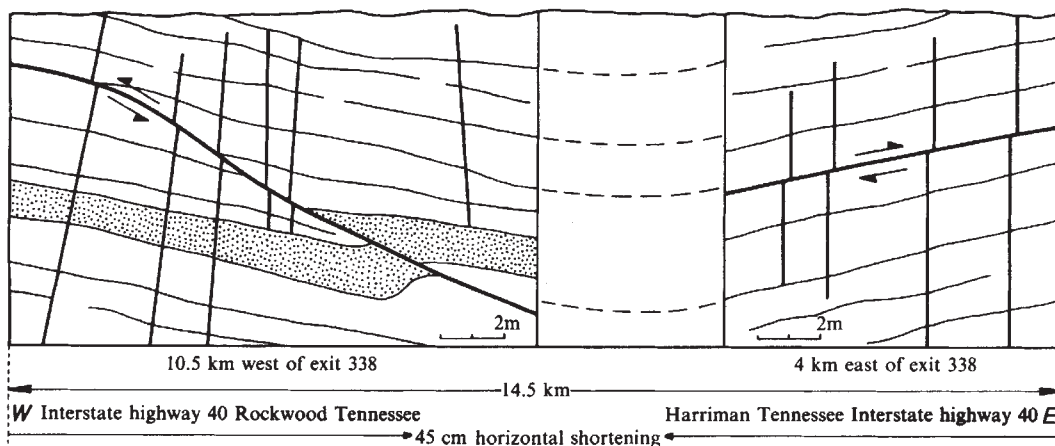


Fig. 4 Recent thrusting and folding of a syncline in Pennsylvanian sandstones and shales (stippled), between Rockwood, Tennessee and Harriman, Tennessee. More or less vertical lines indicate boreholes that are offset at a recently reactivated thrust-fault (left) and along a bedding plane (right).

The rates of relative vertical movements as determined from levelling data of the eastern US reveal that the Appalachians rise relative to the interior platform and to the Atlantic coast at rates of up to 0.6 cm yr^{-1} (ref. 6). This is already two orders of magnitude larger than the estimated average uplift from the early Cretaceous to the present. Brown and Oliver⁶ have published releve profiles with small peaks of increased vertical motion superimposed on the broader uplift of the Appalachians. Two of their profiles traverse close by the sites where recent thrusting was recorded. A peak of prominent uplift at Tyrone, Pennsylvania is located 22 km south of the Port Matilda, Pennsylvania site. Another, but less pronounced peak of vertical motion between Somerset, Kentucky and Harriman, Tennessee is established about 100 km north of the Rockwood-Harriman, Tennessee site in the continuation of the geological strike.

An apparent uplift at the Blue Ridge-Piedmont boundary relative to the Atlantic coastal plain and the Valley and Ridge province is indicated by precise releve along a line from Morristown, Tennessee to Newton, North Carolina and along another traverse in Georgia⁷.

The areas of investigation and generally the central and southern Appalachians are associated with only a moderate seismicity⁸. Nevertheless, the Appalachian mountain belt is contoured by a cluster of epicentres in relation to its foreland and hinterland. The Atlantic coast area is seismically active only in Virginia and South Carolina, the latter area being part of the Charleston-Cumberland seismic zone as an onshore extension of the Atlantic Blake transform fault⁹. The Rockwood-Harriman, Tennessee site of recent thrusting in eastern Tennessee is located at the north-western termination of the Charleston-Cumberland seismic zone. There, the present maximum horizontal north-west-south-east compression ($\sigma_{H_{\max}}$) inferred from thrusting does not correspond to the north-east-south-west-orientation of σ_1 derived from composite fault plane solutions of earthquakes that occurred in the same seismic zone further to the south-east at Lake Jocassee, South Carolina¹⁰ (the seismic data have been revised since instrument polarities were reversed).

The results of structural investigations in the Rockwood-Harriman, Tennessee area contrast also to the *in situ* stress directions of a north-east $\sigma_{H_{\max}}$ and a north-west $\sigma_{H_{\min}}$ obtained from hydrofracturing¹¹. Other $\sigma_{H_{\max}}$ directions determined from hydrofracturing in the western Appalachians from West Virginia to New York State approximately correspond to those derived from recent thrusting at the West Virginia and Pennsylvania sites. Both areas are associated with little seismicity despite prominent recent tectonic events. Some horizontal *in situ* stresses have been determined in the Appalachian Piedmont from Georgia to Pennsylvania by the overcoring technique¹². All sites but one reveal $\sigma_{H_{\max}}$ to be north-east directed. The Appalachians are probably governed by two different stress regimes, one in the west with $\sigma_{H_{\max}}$ orientated north-west to east-west and another in the eastern Appalachians with $\sigma_{H_{\max}}$ predominantly north-east directed. It is assumed that absolute stresses of 50–150 bar as determined by *in situ* overcoring in the Piedmont may be enough to cause recent thrusting like that in the Appalachian Plateau or Valley and Ridge province.

I thank the Departments of Highways at Nashville, Tennessee, Knoxville, Tennessee, and Charleston, West Virginia for dates of the drill periods.

KARLHEINZ SCHÄFER

Institut für Geowissenschaften der Universität,
D-858 Bayreuth, FRG

Received 6 March; accepted 30 May 1979.

1. Rodgers, J. *The Tectonics of the Appalachians* (Wiley, New York, 1970).
2. Cooper, B. N. *Virginia geol. Surv. Bull.* **60** (1944).
3. Gilluly, J. *Bull. geol. Soc. Am.* **75**, 483 (1964).
4. Owens, J. P. in *Studies of Appalachian Geology: Central and Southern* (Cloos Vol.) 417 (Interscience, New York, 1970).
5. Block, J. W., R. C. Clement, L. R. Lew & de Boer, J. *Geology* **7**, 79–81 (1979).

6. Brown, L. D. & Oliver, J. E. *Rev. Geophys. Space Phys.* **14**, 13 (1976).
7. Citron, G. P. & Brown, L. D. *Tectonophysics* **52**, 223 (1979).
8. Bollinger, G. A. *Bull. seism. Soc. Am.* **63**, 1785 (1973).
9. Fletcher, J. B., Sbar, M. L. & Sykes, L. R. *Bull. Geol. Soc. Am.* **89**, 1656 (1978).
10. Talwani, P. *Geophys. pura appl.* **115**, 275 (1977).
11. Haimson, B. C. *ISRM symp.* 23, (Sydney, Australia, Natn. Conf. Publ. 76/4, 1976).
12. Hooker, V. E. & Johnson, C. F. *U.S. Bur. Mines Rep. Inv.* 7224 (1969).

Direct measurements of secondary currents at a river inflexion point

SECONDARY currents are defined as currents which occur in a plane normal to the local axis of primary flow. They may develop in pipes or open channels as a result of non-equal distribution of boundary shear stress or by skewing of cross-stream vorticity into a long stream direction^{1–3}. Secondary currents distort the distributions of primary isovels and boundary shear stress from those found in simple flows and have important effects on sedimentary processes of erosion and deposition^{4,5}. Direct measurements of secondary currents in river bends were reported recently⁶. Further data have been obtained for the reach around the inflexion point between bends (Fig. 1) and are reported here.

At a river bend a strong skew-induced secondary flow cell carries surface water towards the outer bank and bed water towards the inner bank. Measurements made in rivers indicate that a small cell of reverse rotation can exist close to the outer bank in a bend, if that bank is steep⁶.

This cell is generated locally by interaction between the skew-induced cell and the outer bank^{6,7}. It is not connected with the reverse circulation of a relic cell from an upstream bend^{8–10}. Few measurements have been made in the reach between opposite bends, however, to ascertain how the secondary current pattern from the upstream bend decays, to be replaced by that associated with the downstream bend. Some observations have been made in laboratory flumes^{8–10}. Flume data collected by Toebe and Sooky¹⁰ suggest that the skew-induced cell for the downstream bend tends to displace the relic cell laterally towards the outer bank in the downstream bend. Wilson hypothesised that this might be the case on theoretical grounds¹¹. In the flume study the decaying cell does not ever disappear completely, so that a remnant cell persists through the downstream bend.

The secondary flow pattern at the inflexion point shows two approximately symmetrical cells which converge at the surface and diverge at the bed. Chacinski and Francis observed a rather different system in their flume study^{8,9}. Here the skew-induced cell for the second bend originates near the bed and develops upwards, displacing the decaying cell as it does so. This leads to a

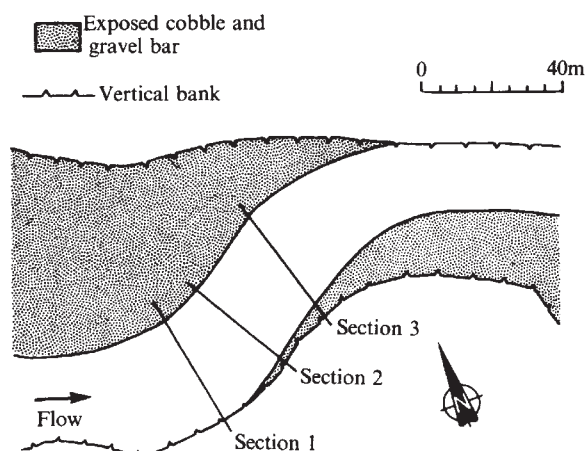


Fig. 1 River Severn at Penstrowed, Powys. Discharge about $13 \text{ m}^3 \text{ s}^{-1}$.

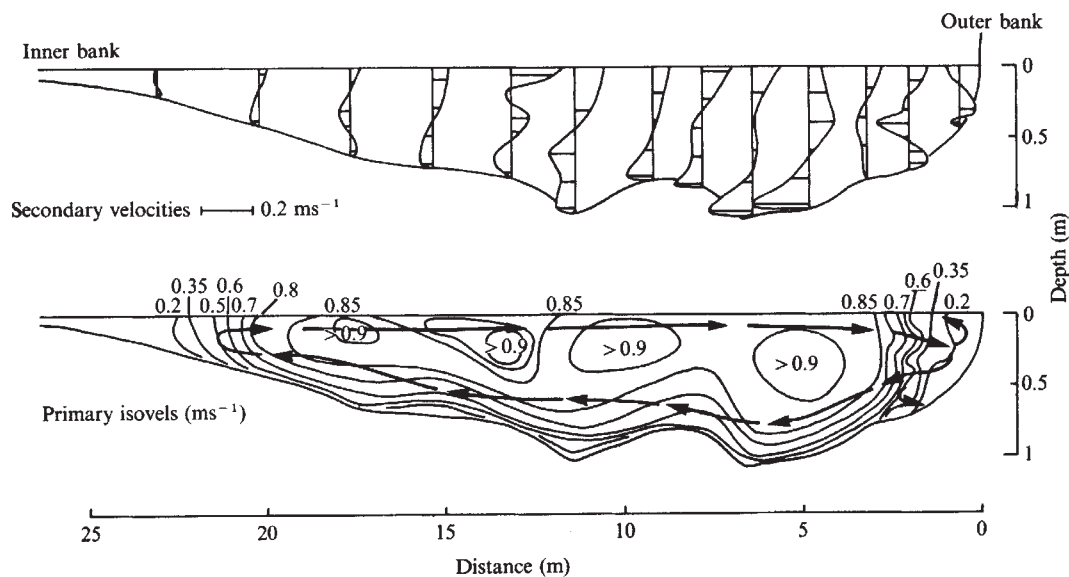


Fig. 2 Primary and secondary velocities at section 1, exit of upstream bend. 19 November 1977, discharge $13 \text{ m}^3 \text{ s}^{-1}$.

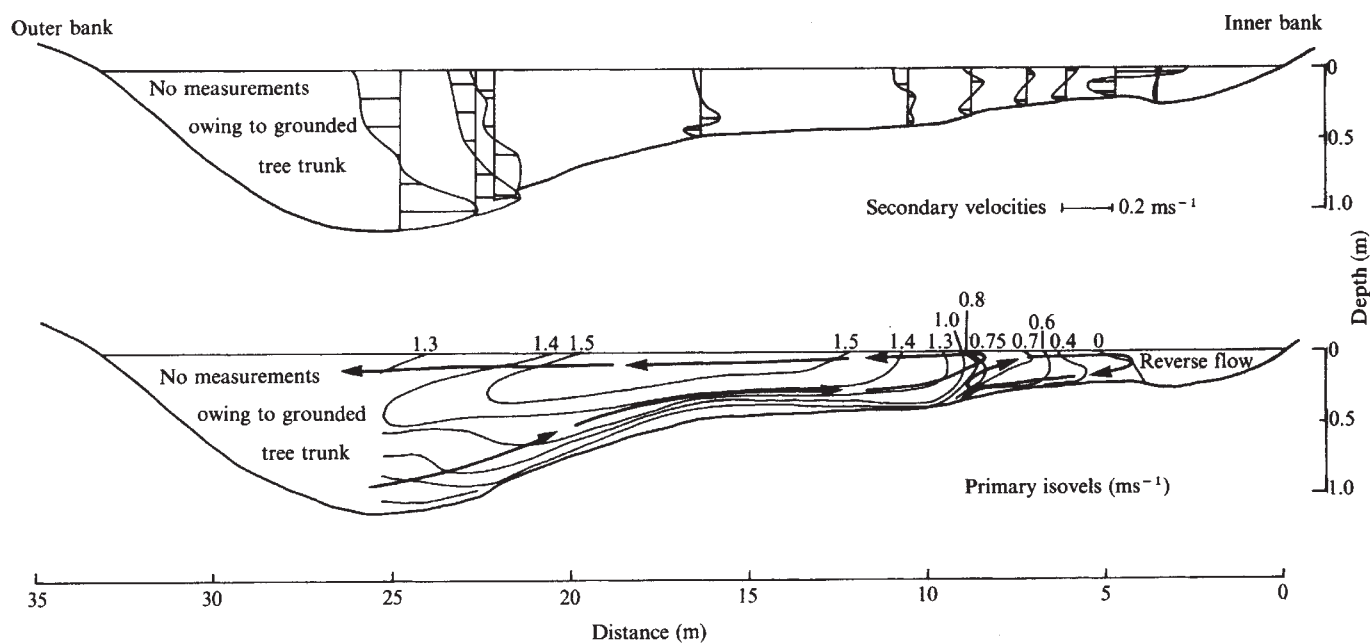


Fig. 3 Primary and secondary velocities at section 3, entrance to downstream bend. 18 November 1977, discharge $17 \text{ m}^3 \text{ s}^{-1}$.

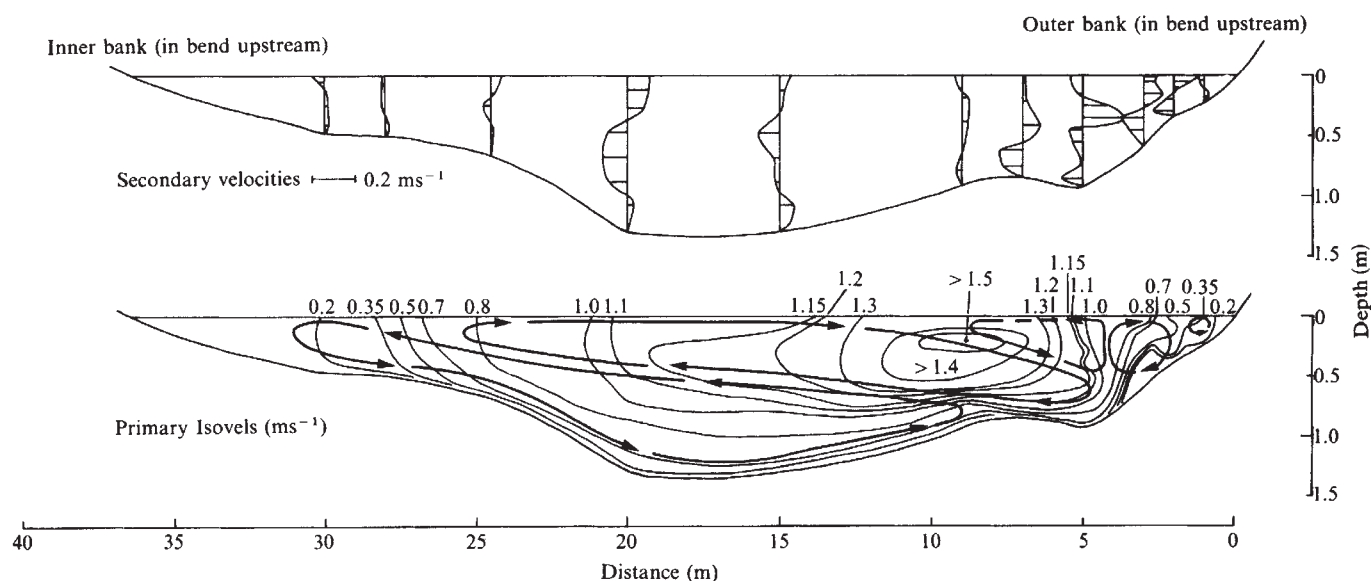


Fig. 4 Primary and secondary velocities at section 2, inflexion point between bends. 17 November 1977, discharge $22 \text{ m}^3 \text{ s}^{-1}$.

system of two stacked cells in the second bend, the upper cell being a remnant of the upstream flow.

The explanation of these apparently contradictory results probably lies in the cross-sectional shape of the flume channels used in the two experiments. In Toebes and Sooky's case the width to depth, or aspect ratio was about 1.5. However, Chacinski and Francis used much higher aspect ratios up to 40 and in some cases 80. As the latter values are more typical of the aspect ratios found in natural streams, it should be expected that the observations of Chacinski and Francis should be more directly applicable to rivers.

Rozovskii¹² used the equations of motion to produce equations describing the development and decay of skew-induced secondary flow. On the basis of these equations, he suggests patterns like those observed by Chacinski and Francis for connecting bends of different directions¹².

To investigate the pattern of secondary flow between bends in a river, data were collected at Penstrowed, Powys, on the River Severn, using an electromagnetic flow meter. This instrument measures velocity in two perpendicular directions simultaneously with an accuracy of 10 mm s^{-1} (refs 13–15). Measurements were made at three sections across the river between two meander bends. The meter was positioned to measure long and cross-stream velocity. These are not necessarily the primary and secondary velocities but a theory previously developed allows us to calculate these components⁶. Although the measurements were made on three consecutive days discharge was not constant. However, the change in discharge did not produce a marked change in the geometry of flow through the reach and so it is still possible to compare qualitatively the patterns of flow at the three sections. Section 1 was located at the exit of the upstream bend, section 2 at the inflexion point, and section 3 at the entrance to the downstream bend (Fig. 1).

The secondary flow pattern at the bend exit (Fig. 2) is dominated by the skew-induced cell from the upstream bend. There seems to be a rather weak outer bank cell near the right hand bank. This is not surprising as secondary currents are known to persist for considerable distances downstream from single bends¹⁶. The data for section 3 (Fig. 3) at the entrance to the downstream bend indicate that this is not the case for a series of bends. Here a skew-induced cell of opposite rotation to that in the upstream bend dominates the secondary flow pattern. Section 2 (Fig. 4) at the inflexion point, seems to show a two cell system with vertical stacking in the central part of the channel. The upper cell is the relic skew-induced cell from the upstream bend, whilst the lower cell is the developing skew-induced cell for the downstream bend. The right-hand margin of the system is complex, but there seem to be outer bank cells resulting from interaction of the skew-induced cells with the bank⁶.

The river measurements are similar to those made by Chacinski and Francis in a laboratory flume of high width-to-depth ratio. A significant difference is that the skew-induced circulation for the downstream bend in the River Severn was fully developed at the bend entrance, whilst in the flume experiments a remnant of the upstream circulation persisted well into the downstream bend. The reason for this difference may lie in the difference in bed topography between the river and the flume. In the river the cross-section at a bend is highly asymmetrical with the deepest point close to the outer bank. At the inflexion point the channel is almost symmetrical, often with a pronounced medial bar. The switching of thalweg from right to left banks between sections 1 and 3 reflects this. It is the curving of streamlines resulting from this effect of bed topography which is initially responsible for the generation of the new skew-induced cell close to the bed. At the bend entrance downstream, plan form curvature reinforces this effect, leading to full development of the skew-induced secondary flow early in the bend. Clearly this would not be the case in the laboratory experiments carried out in flume channels with uniform rectangular cross-sections^{8–10}. In a channel with a uniform cross-section, plan shape curvature alone is responsible for generating the new

skew-induced cell and consequently that cell does not fully displace the relic cell until well into the downstream bend.

Direct measurements of secondary currents at a river inflexion point confirm that the system of cell development proposed by Chacinski and Francis can operate in rivers as well as uniform flume channels. An important difference is that the bed topography in a natural stream seems to result in early and more complete replacement of remnant cells from an upstream bend.

We thank Professor Peter Bradshaw for helpful criticisms.

C R. THORNE
R. D. HEY

*School of Environmental Sciences,
University of East Anglia,
Norwich, UK*

Received 12 April; accepted 14 June 1979.

1. Prandtl, L. *Essentials of Fluid Dynamics* (Blackie, London, 1952).
2. Perkins, H. J. *J. Fluid Mech.* **44**, 721–740 (1970).
3. Bradshaw, P. *An Introduction to Turbulence and its Measurement* (Pergamon, Oxford, 1971).
4. Hooke, R. J. *geol.* **83**, 543–565 (1975).
5. Yen, B. C. *Proc. 16th Congr. IAHR* **2**, 338–346 (1975).
6. Bathurst, J. C., Thorne, C. R. & Hey, R. D. *Nature* **269**, 504–506 (1977).
7. Bathurst, J. C., Thorne, C. R. & Hey, R. D. *J. Hydraul. Div. Am. Soc. civ. Engrs* (in the press).
8. Hey, R. D. & Thorne, C. R. *AREA* **8**, 234–236 (1976).
9. Chacinski, T. M. & Francis, J. R. D. *Trans. Am. geophys. Un.* **33**, 771–773 (1952).
10. Prus-Chacinski, T. M. *Ass. Int. D'Hydrog.* **3**, 311–318 (1954).
11. Toebes, G. H. & Sooky, A. A. *J. Watways Harb. Div. Am. Soc. civ. Engrs* **93**, 213–236 (1967).
12. Wilson, I. G. *Nature* **241**, 393–394 (1973).
13. Rozovskii, I. L. *Acad. Sci. Ukrain. S.S.R. Kiev, USSR* (1957); (translated Prushausky, Y. *Israel Prog. Sci. Trans.* (1961) Monson, S. (Jerusalem), P.S.T. Cat. No. 363 (1963)).
14. Tucker, M. J., Smith, N. D., Pierce, F. E. & Collins, E. P. *J. Inst. Navig.* **23**, 3 (1970).
15. Tucker, M. J. *Proc. Soc. Underwater Technol.* **2**, 174–182 (1975).
16. Kanwisher, J. & Lawson, K. *Limnol. Oceanogr.* **20**, 174–182 (1975).
17. Ippen, A. T. & Drinker, P. A. *J. Hydraul. Div. Am. Soc. civ. Engrs* **88**, 143–179 (1962).

Aneuploidy in mouse fetuses after paternal exposure to X rays

ANEUPLOIDY can result either from non-disjunction, giving trisomy and monosomy, or from chromosome damage leading to loss of chromosomes (monosomy) only. It is known that chromosome loss can be induced after irradiating post-meiotic male germ cells¹, but apart from findings on radiation-induced non-disjunction in oocytes from mice^{2–5}, there is no evidence that paternal exposure to X rays induces chromosomal mal-segregation during male meiosis to produce monosomic and trisomic offspring. Szemere and Chandley⁶ did observe significantly more mal-segregated X and Y chromosomes in spermatocytes II after 100 and 200 rad had been given during preleptotene and pachytene, but failed to recover aneuploid fetuses in progeny from irradiated males. They suggested that only a few aneuploid spermatocytes would be transmitted to the next generation. We have now shown that such induced aneuploidy is indeed transmitted and can be detected in a monosomic and trisomic condition in F_1 fetuses on day 9.5 of gestation.

Twenty 8–10-wk-old random-bred NMRI/Han male mice were given whole-body irradiation of 20 or 200 rad (10 males in each group) and were mated from day 2 to day 36 after irradiation to untreated females of the same age and strain. Non-irradiated males served as matched controls. The radiation source was a Philips-Müller tube with tungsten anode and beryllium window; tube voltage was 150 kV and current 8 mA, Cu filter was 1 mm and the exposure rate was 55 rad min^{-1} . The animals were housed under controlled conditions, darkness lasting from 6 pm to 6 am, tapwater and Altromin pellets (RH 1324) were supplied *ad libitum*. The females were checked every morning (day 0.5) for vaginal plugs and those found to be positive were isolated and kept separately until day 9.5 *post coitum*.

The pregnant females were killed by cervical dislocation, their implantation sites removed and placed into TC 199 culture

Table 1 Aneuploidy in 9.5-d-old mouse fetuses from strain NMRI/Han after paternal irradiation

Dose of X rays given to males (rad)	No. of fetuses (aneuploid/total) karyotyped after G-banding from days					Overall (2-36)
	2-8	9-15	16-22	23-29	30-36	
0						0/211 0.0%
20	0/75	0/33	2/30	0/42	1/36	3/216† 1.4%
200	3/155	0/71	1/20	2/67	1/72	7/385‡ 1.8%* 1.7%*

Matings were from day 2 to day 36 after X-ray exposure. The difference between the control and the irradiated series is statistically significant; there is no significant difference between the 20 rad and 200 rad series.

*Irradiating meiotic and post-meiotic stages $P < 0.05$ (Fisher's exact test).

†Irradiating meiotic stages only, $P < 0.05$ ($F_{2,6}, 0.05 = 5.14 < 5.92$) after 20 rad.

‡Irradiating meiotic stages only, $P < 0.05$ ($F_{2,8}, 0.05 = 4.46 < 5.04$) after 200 rad both compared to the control.

medium containing 17% fetal calf serum, L-glutamine $2 \mu\text{M}$, penicillin $1,000 \text{ IU ml}^{-1}$, streptomycin 10 mg ml^{-1} (reagents from Seromed) and 0.004% colchicine. The embryo proper or the egg membranes only were cultured for 2 h and prepared cytologically according to the method of Evans *et al.*⁷. Chromosomes from at least 7-d-old air-dried slides were stained for G-banding and metaphases were analysed from coded slides. Only one karyotype was prepared from the best 'banded' metaphase with optimal spreading of chromosomes from each embryo or egg membrane. As many extra karyotypes as possible, however, were compared when the first metaphase was abnormal. This procedure of analysis certainly detects those anomalies present in all cells, but would miss some aneuploidies developing only in a mosaic condition.

It has been suggested⁶ that triploid mouse fetuses might arise from irradiation of early spermatocytes. We observed four such fetuses with a karyotype of 60,XXY—three in one litter after 200 rad and one fetus after 20 rad. All triploids were generated, however, from fathers exposed 10 days before mating—after the second meiotic division. This finding, together with the fact that three triploids were derived from one litter and the observation of a high spontaneous background of digynic triploidy in the NMRI/Han mouse strain⁸, suggested that X-ray exposure did not enhance polyploidy in our study.

Overall, 10 of 601 fetuses derived from irradiated fathers were monosomic or trisomic, but no aneuploidy was found among 211 karyotyped fetuses from non-irradiated fathers (Table 1). All the aneuploid fetuses were generated from different fathers and mothers. Our data from the untreated fetuses confirm that aneuploidy is rare in post-implantation mouse embryos not exposed to mutagens.

The identification of the odd chromosome in all 10 aneuploid fetuses provides further information on the survival of monosomics and trisomics during mouse embryogenesis, although chromosomal mosaicism cannot be excluded in our fetuses and in those of other studies⁹. The results shown below are in agreement with the conclusion reached after the studies of robertsonian translocations in the tobacco mouse⁹—that autosomal monosomics are more prone than trisomics to an early elimination during pre- and peri-implantation development.

The following aneuploidies were identified, with the day of mating after irradiation given in parentheses: First group after 20 rad with 41,XX, +11 (21); 41,XY, +2 (21) and 41,XX, +11 (31). Second group after 200 rad with 39,X0 (4); 39,XY, -8 (5); 41,XX, +14 (7); 41,XY, +13 (16); 39,X0 (25); 41,YYY (28); 40,X0, +13 (32). The latter karyotype with 40,X0, +13 is obviously the result of two events: (1) a sex-chromosome loss either by non-disjunction or due to loss after breakage, and (2) due to non-disjunction of chromosome 13. No 41,XXY karyotype was found in our study, in agreement with the observation of only one XXY male among more than 5,000 offspring studied genetically after paternal irradiation¹.

Non-disjunction was induced in mouse oocytes after exposure to equally low doses (10–30 rad) during the preovulatory phase^{2,3}. Aneuploidy was assessed in these studies in oocytes, that is, at a stage long before selection against chromosomally imbalanced fetuses could act. Our observation of 1.4% aneuploid fetuses 9.5 days *post coitum* after paternal exposure to 20 rad would mean that mouse spermatogenesis is at least as sensitive to radiation-induced non-disjunction as oogenesis. Aneuploidy was not significantly enhanced over the effect after 20 rad with the still higher dose of 200 rad. Either some kind of saturation was already reached at the cellular targets for non-disjunction, or the enhanced effects after 200 rad were modified secondarily by, for example, the elimination during spermatogenesis or embryogenesis. [Note that we did not observe any structural anomaly in the 9.5-d-old fetuses after 20 rad, but found balanced and imbalanced translocations after exposure to 200 rad.]

In 6 of the 10 observed aneuploid fetuses the mal-segregation can be allocated with certainty to meiosis I or II. Both primary and secondary non-disjunction were also found to be involved in generating some cases of Down's syndrome^{10,11}. The other four fetuses, however, were derived from sperm irradiated in the vas deferens (39,X0 and 39,XY, -8) or the epididymis (41,XX, +14), or irradiated as mid-spermatids (41,XY, +13). A chromosome loss most probably may explain the X0 karyotype and the monosomy 8, although mitotic non-disjunction during early cleavage cannot be ruled out and should have produced a mosaic condition. This suggestion is not unlikely, because some fetuses with mosaic aneuploidy were found after maternal and paternal irradiation^{5,6}. Mitotic non-disjunction during early cleavages could also explain our two trisomic fetuses (trisomy 13 and 14), generated 7 and 16 d after irradiation. Either we failed to detect the respective monosomic cell lines, or these hypodiploid cells were selectively eliminated earlier—a reasonable suggestion, as monosomics are usually eliminated during the early mouse embryogenesis⁹.

A study of human spontaneous abortions showed that occupational X-ray exposure in man can increase the risk of chromosomally abnormal conceptions¹². Our finding of fetal aneuploidy after paternal irradiation, affecting both autosomes and sex chromosomes, even after exposure to rather low doses, may have implications for the human situation and may lead to a better understanding of the mechanisms involved in non-disjunction.

We thank Mrs H. Schulz, Mrs S. Gebauer and Mr S. Jackson for technical assistance, Mrs B. Schäfer-Berner for photographic work, Dr T. Grimm for advice on the statistical evaluation and Dr P. Virsik for advice on dosimetry. The study was sponsored by a grant of the European Commission 206-76-7 BIO D. Radiation facilities were provided by the Institut für Medizinische Physik und Biophysik (Professor Harder).

I. HANSMANN
R. ZMARSLY
H. D. PROBECK
J. SCHÄFER
J. JENDERNY

*Institut für Humangenetik,
Universität Göttingen,
3400 Göttingen, FRG*

Received 2 April; accepted 29 May 1979.

- Russell, L. B. & Saylor, C. L. *Repair from Genetic Radiation Damage* (ed. Sobels, F. H.) 313 (Pergamon, Oxford, 1953).
- Uchida, I. A. & Lee, C. P. V. *Nature* **250**, 601–602 (1974).
- Uchida, I. A. & Freemant, C. P. V. *Nature* **265**, 186–187 (1977).
- Reichert, W., Hansmann, I. & Röhrborn, G. *Hum. Genet.* **28**, 25–38 (1975).
- Yamamoto, M., Shimada, T., Endo, A. & Watanabe, G. *Nature new Biol.* **244**, 206–208 (1973).
- Szemere, G. & Chandley, A. C. *Mutat. Res.* **33**, 229–238 (1975).
- Evans, E. P., Burtenshaw, M. D. & Ford, C. E. *Stain Technol.* **47**, 229–234 (1972).
- Hansmann, I. & El-Nahass, E. *Cytogenet. Cell Genet.* (in the press).
- Gropp, A. (ed.) *Ciba Fdn Symp.* **40** (1976).
- Langenbeck, U., Hansmann, I., Hinney, B. & Hönig, V. *Hum. Genet.* **33**, 89–102 (1976).
- Jacobs, P. A. & Morton, N. M. *Hum. Hered.* **27**, 59–72 (1977).
- Lazar, P., Gueguen, S. & Boue, A. *Les Accidents Chromosomiques de la Reproduction* (eds Boué, A. & Thibault, C.) 317 (INSERM, Paris, 1973).

Vitamin D and its analogues as a new class of plant growth substances affecting rhizogenesis

SEVERAL substances actively promote adventitious root formation in cuttings: they include indoleacetic acid (IAA), indolebutyric acid (IBA) and α -naphthylacetic acid. Others, which themselves do not or only slightly promote root formation, can act synergistically with one or another of the former to enhance their activity¹. Root formation in cuttings is exploited in the propagation of new varieties of plants, such as polyploids or those with disease resistance. While screening steroids for their effect on rooting in *Populus* spp., we have found a synergistic effect of cortisol on IBA-promoted root initiation². We now report a similar effect of vitamin D and its analogues, and suggest an endogenous function for this group of compounds.

Vitamin D₂ and vitamin D₃ are produced by UV irradiation of the respective provitamins ergosterol and 7-dehydrocholesterol—transformations which can be effected *in vitro* or *in vivo* in animal tissue. The thermal and photochemical pathway involves the sequence: ergosterol, previtamin D₂ and vitamin D₂. Tachysterol and lumisterol, once thought to be intermediates, are now considered to be isomerisation side products³ (Fig. 1).

To test for their effects on plant growth, ergosterol, dihydrotachysterol and vitamin D₂ were applied alone and in the

presence of IBA, to *Populus tremula* (aspen) herbaceous cuttings. Dihydrotachysterol, a reduction product of tachysterol, was used because the latter is not readily available and is highly susceptible to aerial oxidation. Later work (results not given) has shown that dihydrotachysterol and tachysterol have similar activities.

The results (Fig. 2a) show that ergosterol alone has no significant effect at concentrations of 10–50 mg l⁻¹ but that with dihydrotachysterol or vitamin D₂ alone, the number of adventitious roots produced increases by about 60% and 100%, respectively, the maximum effect occurring at 10 mg l⁻¹. IBA alone at 10 mg l⁻¹ promoted rooting as expected, by about 200% (Fig. 2b). Ergosterol in the same concentration range of 10–50 mg l⁻¹, applied with IBA at 10 mg l⁻¹, caused no greater root formation than did the IBA alone, whereas the same concentration range of dihydrotachysterol and vitamin D₂ applied with IBA at 10 mg l⁻¹, caused a pronounced increase in root formation. Thus, in the concentration range 10–50 mg l⁻¹ there was a significant enhancement of root formation (compare this with the similar, but less pronounced, synergistic effect obtained with cortisol and IBA in comparable conditions²) and this effect was dependent on concentration. A similar enhancement of root formation has also been observed with woody cuttings from *Populus tremula* and *Populus nigra* and with cuttings of varieties of *Phaseolus vulgaris*.

Vitamin D₃ has been isolated from *Dactylis glomerata* and various species of grasses⁴ and the presence of vitamin D or its analogues has been suspected in other plant tissues said to have anti-rachitic activity⁵. However, so far no function has been proposed for them in plants. Using the same rooting test, we have found that vitamin D₂ and vitamin D₃ have similar activities (results not shown). Concerning structure/effect relation-

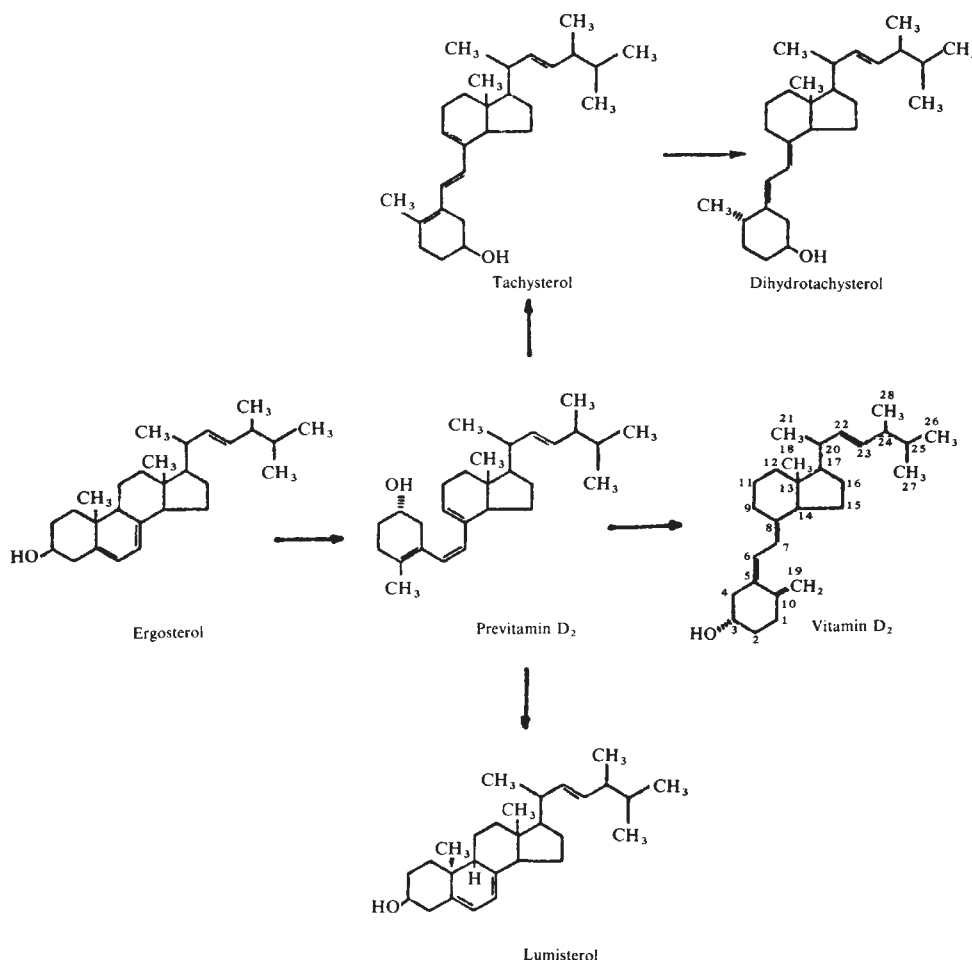


Fig. 1 Thermal and photochemical transformations of ergosterol.

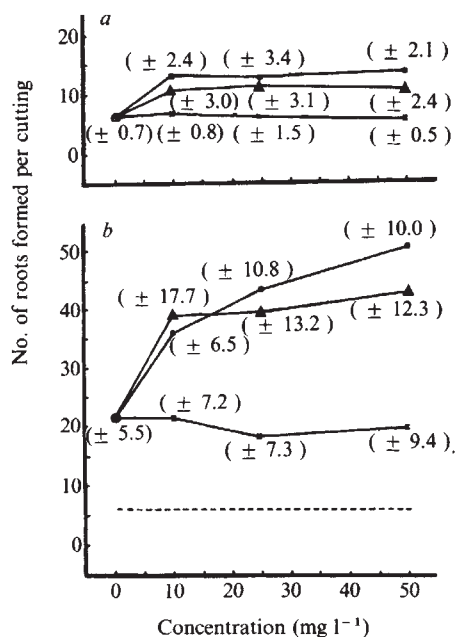


Fig. 2 The *Populus tremula* seedlings were prepared as previously described² and solutions of the test substances were applied to the cuttings for 24 h followed by a change to distilled water. The adventitious roots were counted after 18 d of culture, by which time the number of roots formed was essentially constant, and the results expressed as the mean of at least four culture series (s.d. in parentheses). *a*, Without IBA; *b*, with 10 mg l⁻¹ IBA. ●, Vitamin D₂; ▲, dihydrotachysterol; ×, ergosterol; ---, H₂O control.

ships, we can say that the opening of the steroid ring B of ergosterol to give a seco-steroid is an important step, but changes in the alkyl side chain at C-17 have little or no effect.

A. J. BUCHALA
A. SCHMID

*Institut de Biologie végétale et de Phytochimie,
Université de Fribourg 1700 Fribourg, Switzerland*

Received 3 May; accepted 5 June 1979.

1. Letham, D. S. in *Phytohormones and Related Compounds—A Comprehensive Treatise* Vol. 1 (eds Letham, D. S., Goodwin, P. B. & Higgins, T. J. V.) 349–417 (Elsevier, Amsterdam, 1978).
2. Schmid, A. *C. r. hebdomadaire Séances Acad. Sci., Paris D286*, 1149–1502 (1978).
3. Bell, P. A. in *Vitamin D* (ed. Lawson, D. E. M.) 1–50 (Academic, London, 1978).
4. Raoult, Y., Le Boulch, N., Gounelle, J. C., Marnay-Gulat, C. & Ourisson, G. *Bull. Soc. Chim. biol.* **52**, 641–657 (1970).
5. Grant, A. B. *Nature* **168**, 789 (1951).

Effects of visual experience on the maturation of the efferent system to the corpus callosum

IN normal adult cats the first and second visual areas (V₁ and V₂) in the two cerebral hemispheres are interconnected through the corpus callosum (CC) by neurones (callosal neurones) which have a characteristic laminar location and are restricted to the region of the boundary between cytoarchitectonic areas 17 (V₁) and 18 (V₂)^{1–4}. Callosal neurones acquire this distribution postnatally by disappearing (dying, or eliminating their callosal connections) from most of areas 17 and 18 and decreasing in number in layer VI (refs 5, 6). Here we show that although kittens deprived of pattern vision by binocular lid suture still develop an apparently normal distribution of callosal neurones, monocularly deprived or enucleated, or bilaterally squinted kittens retain to adulthood some callosal neurones in parts of

area 17 where such neurones are not normally found. Thus, visual experience can influence the selection of callosal neurones although this selection is probably influenced by more than one factor.

The distribution of callosal neurones in 9 abnormally reared (Table 1) and 14 normal control^{1,4} cats was determined by the retrograde transport of horseradish peroxidase (HRP) injected at 10–12 sites within the contralateral visual cortex. One month

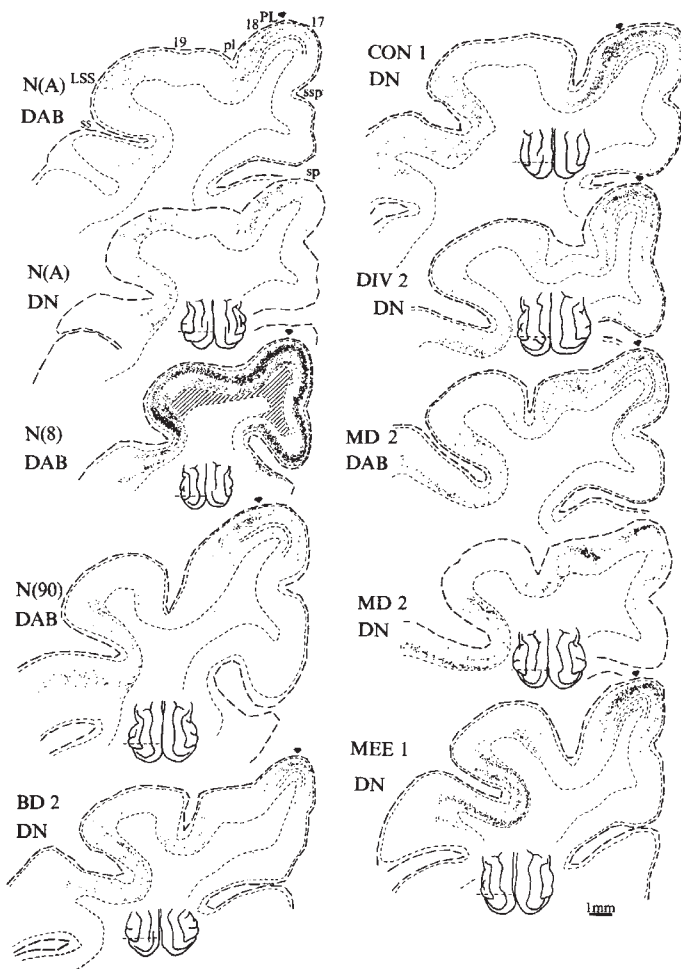


Fig. 1 Computer microscope²⁶ charts of the distribution of HRP-labelled callosal neurones in 80-μm thick coronal sections at comparable rostrocaudal levels of the visual cortex in different animals. Multiple, 0.5-μl injections of HRP (Boehringer grade I, 33% in sterile saline) were placed in the contralateral cortex. After a 2-day survival time, adjacent, 1-in-4 series of 80-μm sections of each brain were processed with diaminobenzidine (DAB)^{1,4} and *O*-dianisidine (DN)²⁷. Adjacent sections were reacted with the two substrates as shown for the normal adult, N(A), and for MD 2. In all animals, the injected postlateral and lateral gyri (areas 17, 18, 19) were invaded almost completely by dark HRP precipitate; HRP-filled neurones were found over most of the mediolateral extent of the ipsilateral lateral geniculate nucleus. Each HRP-labelled neurone is represented by a black dot. The level of each section is indicated on the inset drawn from a dorsal view photograph of the corresponding brain. In each section the thin, interrupted lines mark the lower boundaries of layers I, IV, V and VI (in N(8)), lower boundaries of layers IV and V are not shown). In the DN sections from N(A) and MD 2 only the lower boundary of layer VI is marked. Cross hatching in N(8) indicates a region of displaced (migrating?) neuronal somata. The two kittens N(8) and N(90) come from a different study^{5,6}. Their ages at the time of HRP injection were 8 and 90 days, respectively. Abbreviations: ss, pl, ssp and sp: suprasylvian, postlateral, suprasplenial and splenial sulci, respectively. PL: postlateral gyrus. LSS, 19, 18 and 17 indicate the lateral suprasylvian visual areas, and areas 19, 18 and 17, respectively. The arrow indicates the boundary between the latter two areas.

before death, the strabismic animals could visually guide their paws with either eye; single unit recording from their areas 17 and 18 between the time of HRP injection and death showed that unlike normal cats, but as in monocularly squinted cats⁷, most cortical neurones had monocular receptive fields.

In normal adult cats, the cortical volume containing callosal

neurones (callosal efferent zone; CZ) consists of two radially separated laminae, located in layers III/IV and VI (subzones a and c, respectively). Subzone a is wider and richer in callosal neurones. Its medial boundary can extend to the crest of the lateral-postlateral gyrus in area 17 (Figs 1, 2, N(A)) and occasionally, at rostralmost levels, as far as the suprasplenial

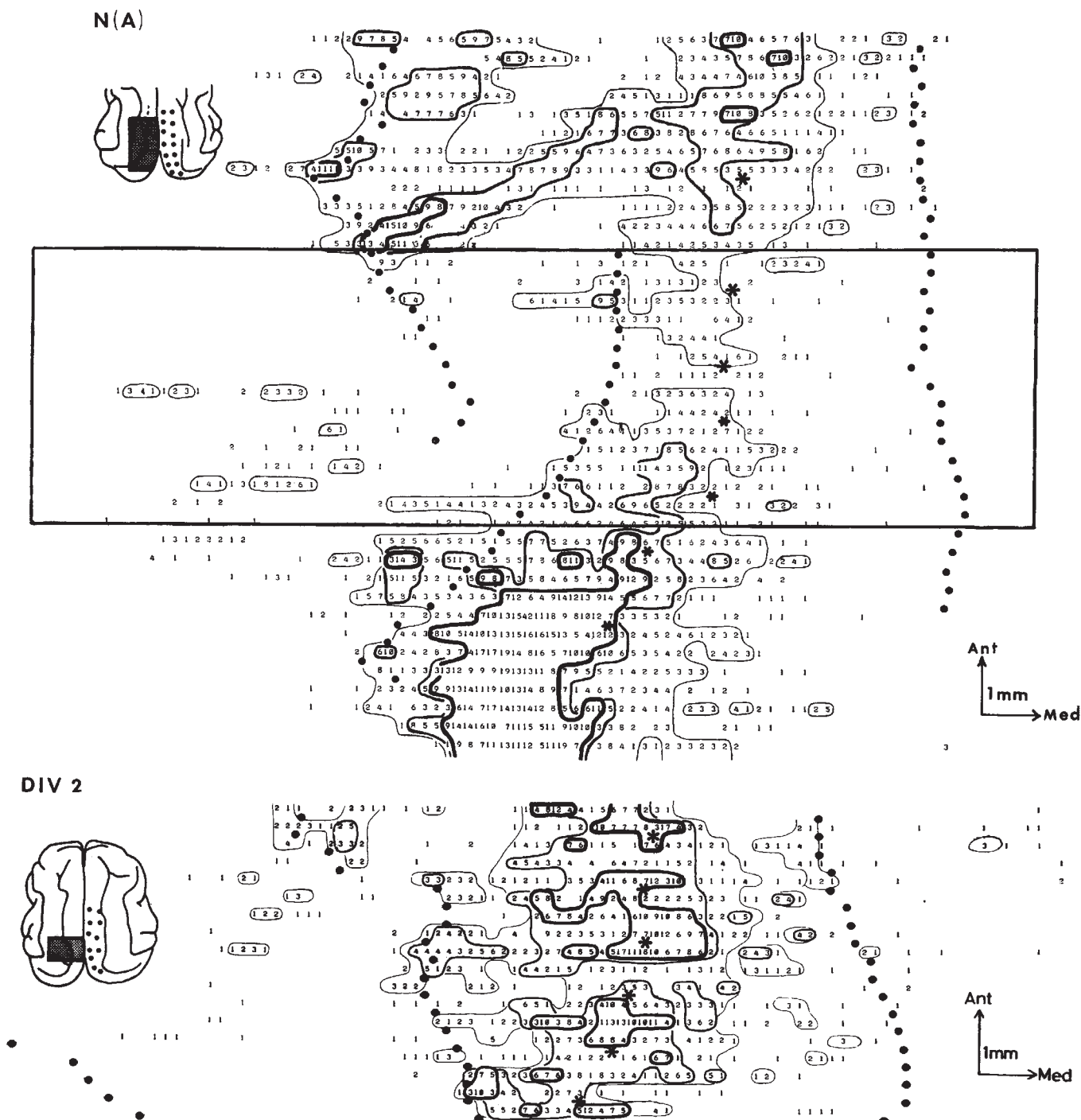


Fig. 2 Computer reconstructions of part of callosal subzone a in a normal adult cat (N(A)) and in a cat raised with divergent strabismus (DIV 2), processed with *O*-dianisidine. Flattened views of the postlateral and lateral gyri (stippled area in dorsal view insets). Dotted lines indicate, from left to right, the fundi of the lateral, postlateral and suprasplenial sulci. A rarely seen branch of the postlateral sulcus was found in DIV 2. The asterisks mark the boundary of areas 17 and 18. The part of the reconstruction of brain N(A) within the rectangle corresponds in rostrocaudal coordinates to the part reconstructed for DIV 2. The neurones in layers II-IV of each section were projected onto a line running parallel to the pial surface and 400 μ m deep; the line was then divided into bins of 200 μ m. Each horizontal row of numbers in the reconstruction represents the counts of labelled cells for each bin of one section. Contour lines of increasing thickness are drawn around regions where the summed count for any two adjacent bins is 1, 2 and 4 times, respectively, the average count for individual bins (all bins containing at least one labelled neurone were used to obtain this average). The sections were spaced 320 μ m apart in N(A) and 280 μ m apart in DIV 2. Because of this, the scales of the rostro-caudal axes in the two reconstructions are different. Sites of injection are marked by dots on the insets. For each animal, the homogeneity of the injection site was judged from that of retrograde labelling in the ipsilateral lateral geniculate body.

Table 1 Experiments on the effects of abnormal visual experience on the development of the CZ

Animal	Age at first operation and at death‡	Angle of squint (deg)
BD 1	7; 128	
BD 2	6; 141	
MD 1*	8; 173	
MD 2*	8; 146	
CON 1†	8; 250	26
CON 2†	9; 246	39
DIV 1†	9; 288	33
DIV 2	6; 349	45
MEE 1*	1; 124	

BD, Binocular lid suture; MD, monocular lid suture; CON, bilateral section of lateral rectus muscles (convergent strabismus); DIV, bilateral section of medial rectus muscles (divergent strabismus); MEE, monocular enucleation. Angle of squint was determined by ophthalmoscopic projection of retinal landmarks in paralysed animals (in normal animals about 3.50 (ref. 25)).

* Injected contralateral to open (remaining) eye.

† Electrophysiological experiment.

‡ Day of birth = day 0; eye opening normally around day 8.

sulcus; its lateral boundary is irregular and susceptible to individual variation. However, callosal neurones are generally found along the medial banks and sometimes down to the fundi of the lateral and postlateral sulci.

At the ages at which experimental rearing was started, callosal neurones are found throughout areas 17 and 18 (Fig. 1, N(8)) and (especially in subzone c) are more densely packed than in the adult. In normal animals the callosal neurone distribution is apparently mature at 3 months (Fig. 1, N(90)), 34 days before our youngest kitten was killed.

In the binocularly deprived cats (Fig. 1, BD 2) the medio-lateral and radial distributions of callosal neurones seem to be very similar to those of normal adults. The CZ may be slightly narrower and callosal neurones a little less densely packed than in normal cats; additional quantitative data are necessary to decide this point.

In all other experimental animals (Fig. 1, right-hand column), subzone a extends further medially in area 17 than in binocularly deprived and normal animals. Callosal neurones are found along the medial bank of area 17 and reach the fundus of the suprasplenic sulcus along almost all its rostrocaudal extent. In CON 1 and DIV 2 a few neurones are also found in the part of area 17 between the suprasplenic and splenic sulci and in the dorsal bank of the latter (see Figs. 1, 2). Because of the normal irregularity and variability of the lateral border of the CZ, it is difficult to determine whether the CZ is also wider than normal in area 18. It is not narrower, so that the presence of callosal neurones in the medial part of area 17 cannot be ascribed to an overall medialward displacement of the CZ.

In the monocularly enucleated cat (Fig. 1, MEE 1) subzone c contains more callosal neurones than in normal or in the other experimental animals.

These results indicate that visual experience can affect which juvenile callosal neurones (or axons) are stabilised and/or which are eliminated, although it is possible that it induces formerly non-callosal neurones to produce callosal axons. The present limited effects suggest either that we did not find the most effective paradigm of visual deprivation or, more likely, that information of non-visual origin is also used in the development of the CZ. There seem to be two possible sources for such information: (1) the topographic (or retinotopic) organisation of structures projecting to or receiving from the cortex containing the callosal neurones; (2) an innate programme within the callosal neurones themselves.

The abnormal distribution of callosal neurones and terminals in Siamese cats^{2,8} seems to support the first possibility. However,

the development of callosal connections in binocularly deprived Siamese cats must be investigated to differentiate the effects of genetic abnormalities in the retino-geniculo-cortical topology from those of strabismus, which also affects this genotype. The mechanisms, time course and site of action of factors influencing the development of the CZ remain to be determined. The enlargement of the CZ may be related to the almost complete loss of binocular responsiveness of cortical neurones provoked by strabismus⁷, monocular deprivation⁹ or enucleation.

The reduction of the CZ is probably one aspect of a more general ontogenetic diminution of neurones and/or neuronal connections (for examples see refs 10–13). Our results and those obtained in Siamese cats^{2,8}, chicks¹⁴ and monkeys^{13,15} suggest that transitory neonatal connections can provide a substrate for the plasticity of neuronal connectivity. However, after transitory callosal neurones have been eliminated modifications of callosal terminals may still occur¹⁶.

One functional consequence of the enlarged CZ is related to the fact that the tangential position of callosal neurones determines their receptive field location. Thus, the receptive fields of callosal neurones in the expanded CZ probably include an abnormally large portion of the visual field. However, they should still be concentrated near the vertical meridian^{8,17–19}, represented at the 17/18 border^{20,21}.

Callosal neurones are thought to integrate the cortical representation of the visual field and subserve binocular interactions contributing to stereopsis and possibly eye movements (for reviews see refs 22, 23). In the amphibian *Xenopus laevis*, binocular vision seems to be subserved by intertectal connections whose development is affected by visual experience²⁴, permitting compensation for normal or experimentally induced changes in ocular alignment. It is attractive to view the modifications we observe in the cat CZ as partial compensation for the disturbance of the binocular visual field, although the adaptiveness of the changes remains to be demonstrated.

This work was supported by Swiss NSF grant 3.319.0 78 to G.M.I., D.O.F. received a fellowship from the American-Swiss Federation for Scientific Exchange.

Note added in proof: Three more kittens deprived of vision by bilateral eyelid closure for 1–4 months after birth have now been examined. In all these animals MRP-labelled callosal neurones were fewer than in normal cats and showed radial and tangential distributions similar to that of binocularly deprived animals described above. Quantitative analysis of the reduction in number of callosal neurones following binocular deprivation is under way.

G. M. INNOCENTI
D. O. FROST

*Institute of Anatomy,
University of Lausanne,
Rue du Bugnon 9,
CH-1011 Lausanne-CHUV, Switzerland*

Received 15 March; accepted 24 May 1979.

- Innocenti, G. M. & Fiore L. *Neurosci. Lett.* **2**, 245–252 (1976).
- Shatz, C. J. *J. comp. Neurol.* **173**, 497–518 (1977).
- Sanides, D. & Donato-Oliver, F. in *Architectonics of the Cerebral Cortex* (eds Brazier, M. A. B. & Petsche, H.) 227–234 (Raven, New York, 1978).
- Innocenti, G. M. (in preparation).
- Innocenti, G. M., Fiore, L. & Caminiti, R. *Neurosci. Lett.* **4**, 237–242 (1977).
- Innocenti, G. M. *Archs ital. Biol.* **116**, 463–470 (1978).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **28**, 1041–1059 (1965).
- Shatz, C. J. *comp. Neurol.* **171**, 229–246 (1977).
- Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **28**, 1029–1040 (1965).
- Hamburger, V. & Levi-Montalcini, R. *J. exp. Zool.* **111**, 457–501 (1949).
- Redfern, P. A. *J. Physiol., Lond.* **209**, 701–709 (1970).
- Clarke, P. G. H., Rogers, L. A. & Cowan, W. M. *J. comp. Neurol.* **167**, 125–142 (1976).
- Rakic, P. *Nature* **261**, 467–471 (1976).
- Hollyday, M. & Hamburger, V. *J. comp. Neurol.* **170**, 311–320 (1976).
- Hubel, D. H., Wiesel, T. N. & Le Vay, S. *Phil. Trans. R. Soc. B* **278**, 377–409 (1977).
- Lund, R. D., Mitchell, D. E. & Henry, G. H. *Brain Res.* **144**, 169–172 (1978).
- Choudhury, B. P., Whitteridge, D. & Wilson, M. E. *Q. J. exp. Physiol.* **50**, 214–219 (1965).
- Berlucchi, G., Gazzaniga, M. S. & Rizzolatti, G. *Archs ital. Biol.* **105**, 583–596 (1967).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **30**, 1561–1573 (1967).

20. Bilge, M., Bingle, A., Seneviratne, K. N. & Whitteridge, D. J. *Physiol., Lond.* **191**, 116P (1967).
 21. Tusa, R. J., Palmer, L. A. & Rosenquist, A. C. *J. comp. Neurol.* **177**, 213–236 (1978).
 22. Berlucchi, G. *Brain Res.* **37**, 371–392 (1972).
 23. Doty, R. W. & Negrao, N. in *Handbook of Sensory Physiology*, Vol. 73 (ed. Jung, R.) 543–582, Springer, Berlin (1973).
 24. Keating, M. J. *Phil. Trans. R. Soc. Lond.* **B278**, 277–294 (1977).
 25. Vakkur, G. J., Bishop, P. O. & Kozak, W. *Vision Res.* **3**, 289–314 (1963).
 26. Glaser, E. M. & Van der Loos, H. *IEEE Trans. biomed. Engng* **12**, 22–31 (1965).
 27. Coleman, D. R., Scalia, F. & Cabrales, E. *Brain Res.* **102**, 156–163 (1976).

Electroconvulsive shock treatment decreases β -adrenergic receptor sensitivity in rat brain

ELECTROCONVULSIVE SHOCK (ECS) is one of the most effective forms of therapy available for patients with depression¹. Its mechanism of action, however, is not clear. Like other antidepressants, ECS probably augments the postsynaptic availability of neurotransmitters, for it increases the release and turnover of noradrenaline (NA)^{2–4}, decreases NA uptake⁵, and raises the activity of tyrosine hydroxylase. Also, synthesis of cyclic AMP stimulated by NA (NA-sensitive adenylate cyclase) is reduced in brain tissues obtained after chronic treatment with ECS^{6,7}. It has been reported that chronic administration of antidepressant drugs reduces β -adrenergic receptor density in the rat brain⁸. Because ECS is another form of treatment for depressive illness, we examined whether chronic administration of ECS has effects similar to those of other antidepressants on β -adrenergic receptor sensitivity in rat brain. We report here that we observed a significant decrease in β -adrenergic sensitivity following chronic ECS treatment. The relationship between such observed effects and clinical response requires further investigation.

Binding studies with radiolabelled ligands like ³H-dihydroalprenolol (³H-DHA) or [¹²⁵I]-hydroxybenzylpindolol (β -adrenergic antagonists) have been used for identification and characterisation of β -adrenergic receptors in the brain^{9–11} and other tissues¹²; thus, they provide a method for direct quantitation and study of the kinetic properties of these receptors.

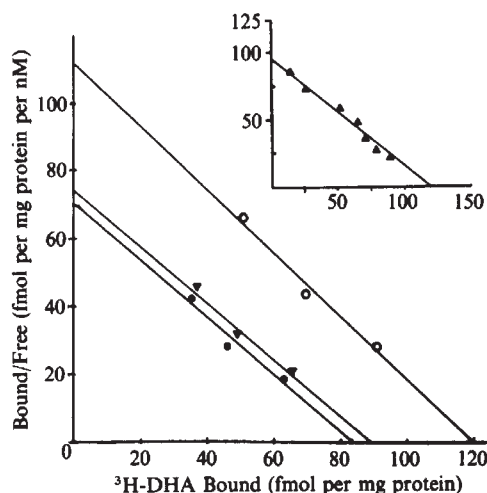


Fig. 1 Scatchard analysis of specific ³H-DHA binding to the brain particulate fractions obtained from control rats (○), DI-treated rats (●) and ECS-treated rats (▲). The experimental conditions are described in Table 1 and each point is the mean obtained from six animals. Receptor density (B_{max}) and dissociation constant (K_D) were calculated from this figure and are shown in Table 1. Inset: Experimental conditions are described in Table 1. The specific ³H-DHA binding to particulate fractions was determined as a function of increasing ³H-DHA concentrations (0–4.5 nM) in a control rat. Scatchard analysis of the obtained data as shown in the inset gave a B_{max} of 115 fmol ³H-DHA bound per mg protein and an apparent K_D of 1.12 nM. This experiment is one of several experiments carried out in identical conditions and giving similar results.

Table 1 Effect of chronic treatment with desipramine and electroconvulsive shock on receptor density and dissociation constant of β -adrenergic receptors in rat brain homogenates

Treatment	Receptor density (fmol per mg protein)	% of control	P values as compared with control	Dissociation constant (nM)
Control (n = 6)	119 ± 8.8			1.06 ± 0.05
Desipramine (n = 6)	84 ± 3.5	71	<0.001	1.19 ± 0.06
ECS (n = 6)	87 ± 3.5	73	<0.001	1.14 ± 0.06

Male Sprague–Dawley rats (100–150 g) were subjected to ECS (75 mA for 1 s through transcorneal electrodes) once daily for 12 d. Another group was injected with desipramine (10 mg per kg, intraperitoneally) once daily for the same period. Control rats were injected with normal saline and were handled similarly to the ECS rats, but did not receive shocks. The rats were killed 1 h after the last treatment, and the cortices were removed, homogenised in 50 volumes (w/v) of Tris buffer (0.05 M, pH 7.4), and spun at 49,000g for 10 min. This process was repeated, and the pellet was finally suspended in Tris buffer. Binding of ³H-DHA was carried out in triplicate by incubation of an aliquot of the suspension (0.3–0.5 mg protein) with ³H-DHA at three concentrations (between 0.4 and 4 nM) in Tris buffer in a final volume of 1 ml, with or without 1 μ M (–)propranolol, at 23 °C for 15 min. ³H-DHA bound to the membrane was separated by filtration through a Whatman (GF/F) filter and washed three times with 4 ml buffer. The filter was placed in a BBS cocktail and counted in a Nuclear Chicago scintillation spectrometer. Specific binding is defined as the difference between total binding and the binding observed in the presence of 1 μ M (–)propranolol. Receptor density, as shown in Fig. 1, was calculated by Scatchard analysis, and is the intercept on the x axis (B_{max}). K_D values were calculated by Scatchard analysis; P values were obtained by paired t-test (two tail).

We therefore measured the binding ³H-DHA to particulate fractions prepared from rat cerebral cortex. The method used was essentially that of Bylund and Snyder⁹, and we obtained binding characteristics similar to those they reported. The binding was rapid, stereospecific to β -adrenergic agonists and antagonists and showed saturation kinetics similar to those reported. Scatchard analysis of the ³H-DHA binding to rat cerebral cortex resulted in a straight line, indicating a single class of high-affinity receptors (Fig. 1, inset). The maximum number of binding sites (receptor density) observed (115 fmol per mg protein) and the observed apparent dissociation constant (1.12 nM) were also similar to those previously reported^{8,9}.

The effects of chronic treatment with ECS and desipramine (DI) on β -adrenergic receptor sensitivity (Table 1) were assessed by determining the maximum number of ³H-DHA binding sites by Scatchard analyses (Fig. 1). Binding of ³H-DHA to particulate fractions obtained from the cortex was carried out at three different concentrations of ³H-DHA. The observed mean β -adrenergic receptor density for the control rats was 119 ± 8.8 (fmol ³H-DHA bound per mg protein). DI caused a 29% reduction in receptor density ($P < 0.001$), in agreement with the results reported by Banerjee *et al.*⁸. Treatment with ECS resulted in a 27% decrease ($P < 0.001$) in density. The mean apparent dissociation constants (K_D) of ³H-DHA binding, determined by Scatchard analysis (Fig. 1), were similar for the treated and control rats (Table 1), indicating that the receptor density changes were not related to alterations in the affinity of receptors.

These results thus indicate that chronic ECS treatment produces a reduction in the sensitivity of β -adrenergic receptors in rat brain. This effect is similar to the effects of other antidepressants, for example, DI, presented here, and DI, doxepin and iprindole, reported by Banerjee *et al.*⁸. The results also suggest that a decrease in NA-stimulated synthesis of cyclic AMP in rats chronically treated with ECS, as reported previously^{6,7}, may be related to a decrease in β -adrenergic receptor sensitivity.

Treatment with ECS causes increased release and turnover of NA^{4,5,13} and possibly of dopamine in various areas of the rat brain. Tricyclic antidepressants, on the other hand, inhibit reuptake of biogenic amines in synaptosomal preparations¹⁴, and drugs like pargyline and phenelzine inhibit monoamine oxidase activity. These antidepressants have different modes of action

on the biogenic amine system; however, all increase the biogenic amine concentration in the postsynaptic area, and reduce the sensitivity of β -adrenergic receptors⁸ and the responsiveness of adenylate cyclase to NA, as reported by others^{6,7,15,16} and presented here. Furthermore, Deguchi and Axelrod¹⁷ and others¹² have suggested that continued exposure of adenylate-cyclase-coupled β -receptors to elevated levels of agonists lowers the sensitivity of the receptors; we might therefore assume that loss of β -adrenergic receptor sensitivity on chronic treatment with antidepressants is caused by overexposure of postsynaptic β -adrenergic receptors to elevated levels of NA. However, iprindole has no significant effect on re-uptake, release or turnover of catecholamines¹⁸, and it is not certain whether it increases the levels of NA in the synaptic cleft; like other antidepressants, it reduces β -adrenergic receptor sensitivity⁸ and decreases NA responsiveness of adenylate cyclase¹⁶. Therefore, as suggested by Banerjee *et al.*⁸, other mechanisms may also be involved in the development of the reduction in β -adrenergic receptor sensitivity.

The clinical effects of antidepressants are observed only after 1–3 weeks of treatment. As the effects of antidepressants on β -adrenergic receptor sensitivity are observed after a similar treatment period, the two may be closely related. However, the time course of development of reduced β -adrenergic receptor sensitivity must be studied before such a relationship can be confirmed. It is also possible that the effects of antidepressants on β -adrenergic receptor sensitivity are secondary to their effects on the catecholamine systems.

GHANSHYAM N. PANDEY
WILLIAM J. HEINZE
BARBARA D. BROWN
JOHN M. DAVIS

Research Department,
Illinois State Psychiatric Institute,
1601 W. Taylor Street,
Chicago, Illinois 60612

Received 28 February; accepted 22 May 1979.

1. Avery, D. & Winokur, G. *Biol. Psychiat.* **12**, 507–523 (1977).
2. Ebert, M. H., Baldessarini, R. J., Lipinsky, J. F. & Berv, K. *Archs gen. Psychiat.* **29**, 397–401 (1973).
3. Modigh, K. *Psychopharmacology* **49**, 179–185 (1976).
4. Thiery, A. M., Javoy, F., Glowinsky, J. & Kety, S. S. *J. Pharmac. exp. Ther.* **163**, 163–171 (1968).
5. Hendley, E. P. & Welch, B. L. *Life Sci.* **16**, 145–154 (1975).
6. Pandey, G. N. & Davis, J. M. in *Neuropharmacology of Cyclic Nucleotides* (ed. Palmer, G. C.) (Urban and Schwarzenberg, Baltimore, in the press).
7. Vetulani, J., Stawarz, R. J., Dingell, J. V. & Sulser, F. *Naunyn-Schmiedeberg's Archs Pharmac.* **293**, 109–114 (1976).
8. Banerjee, S. P., Kung, L. S., Riggi, S. J. & Chanda, S. K. *Nature* **271**, 380–381 (1977).
9. Bylund, D. B. & Snyder, S. M. *Molec. Pharmac.* **12**, 568–580 (1976).
10. Bylund, D. B., Charness, M. E. & Snyder, S. H. *J. Pharmac. exp. Ther.* **201**, 644–653 (1977).
11. Alexander, R. W., Davis, J. N. & Lefkowitz, R. J. *Nature* **258**, 437–439 (1975).
12. Mukherjee, C., Caron, M. G. & Lefkowitz, R. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1965–1969 (1975).
13. Musacchio, J. M., Julou, L., Kety, S. S. & Glowinsky, J. *Proc. natn. Acad. Sci. U.S.A.* **63**, 1117–1119 (1969).
14. Glowinsky, J. & Axelrod, J. *Nature* **204**, 1318–1319 (1964).
15. Frazer, A. *et al. Neuropharmacology* **2**, 211–220 (1974).
16. Vetulani, J. & Sulser, F. *Nature* **257**, 495–496 (1975).
17. Deguchi, T. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2411–2414 (1973).
18. Rosloff, B. N. & Davis, J. M. *Psychopharmacologia* **40**, 53–64 (1974).

How does adrenaline accelerate the heart?

THE way in which adrenaline acts on the sinoatrial (SA) node to accelerate the heart rate has hitherto been obscure. However, in various other parts of the heart adrenaline increases the slow inward ($\text{Ca}^{2+}/\text{Na}^{+}$) current^{1–4}, and voltage-recording experiments have indicated that adrenaline also has this action in the sinus region^{5–7}. In the voltage-clamp experiments reported here, we find that adrenaline does indeed increase the slow inward current in the SA node of the rabbit, but that it also augments the

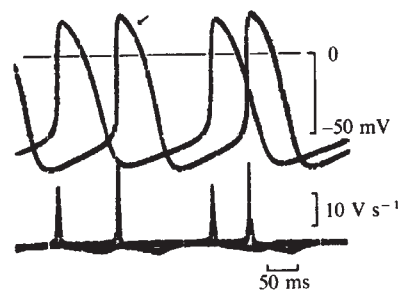


Fig. 1 Intracellular record of spontaneous activity in the rabbit SA node before and during (trace indicated by arrow) perfusion with adrenaline, 5×10^{-8} M. Lower trace shows the rate of change of voltage. In this and in the following figures traces have been recorded on tape during the experiment and overlapped on a storage oscilloscope on replay.

outward current which would tend to decelerate pacemaker depolarisation. We find that an additional current, i_t , is activated within the range of voltage where the pacemaker depolarisation occurs: this could be important both in normal pacemaking and in adrenaline-induced acceleration.

Figure 1 shows the effect of adrenaline on the intracellular voltage record and the rate of change of voltage (dV/dt) in a small, spontaneously firing preparation (about $0.3 \text{ mm} \times 0.3 \text{ mm}$) from the rabbit SA node. The overshoot of the action potential increases, as does dV/dt during the upstroke. As any fast sodium current present was blocked by the presence of tetrodotoxin ($5 \times 10^{-7} \text{ g ml}^{-1}$) these effects are probably caused by an increase in the slow inward current⁸. Figure 1 also indicates that the adrenaline produces a slightly more negative maximum diastolic potential, and an acceleration of the rate of repolarisation, both of which imply an increase of total membrane current in the outward, rather than the inward direction.

Voltage clamp of SA nodal tissue is much more difficult than that of other cardiac tissues; only recently has a method, pioneered by Noma and Irisawa¹³, been developed for successfully clamping the mammalian sinus. In this, very small ($0.3 \text{ mm} \times 0.3 \text{ mm}$) preparations which are still spontaneously active are dissected from the SA node and voltage clamped using two microelectrodes. Good voltage-clamp results can be obtained over a range of potentials, including the pacemaker range, although in response to large depolarisations and hyperpolarisations voltage control is often lost.

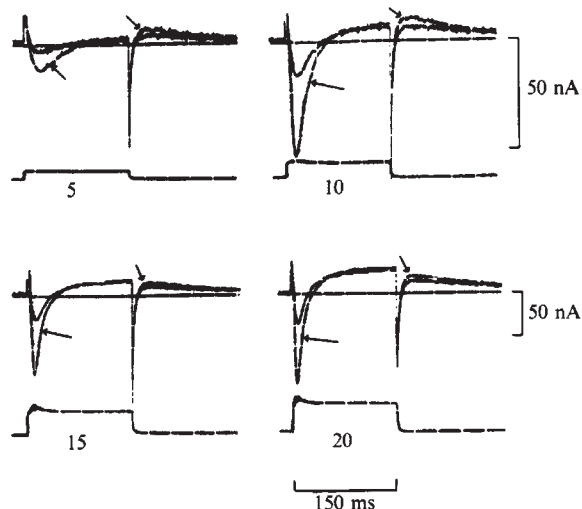


Fig. 2 Increase of $i_{\text{Ca/Na}}$ and i_{K} in adrenaline. Voltage-clamp depolarisations of 5, 10, 15 and 20 mV from a holding potential of -45 mV (lower traces in each panel) were given before and during the action of adrenaline, 5×10^{-7} M. All current traces recorded in adrenaline (indicated by arrows) have been shifted upwards by 2 nA. Solid lines represent zero current.

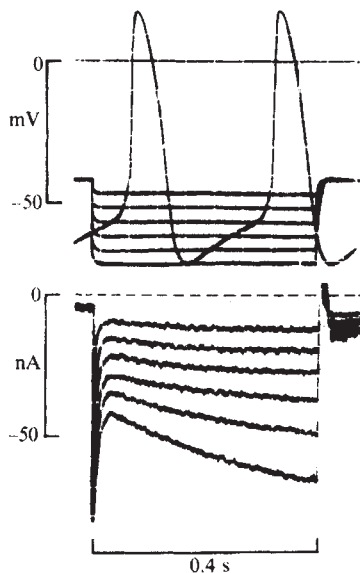


Fig. 3 *a*, Superimposition of pacemaker activity and voltage-clamp hyperpolarisations from a holding potential of -42 mV in the same preparation; *b*, currents recorded during the clamp hyperpolarisations. Note that the time-dependent changes in i_t are greater at more hyperpolarised levels.

With this method it can be shown that in the rabbit SA node adrenaline does indeed increase both slow inward current ($i_{Ca/Na}$) and outward potassium current (the current termed i_K by DiFrancesco Noma and Trautwein¹⁰). Figure 2 shows the increase in peak $i_{Ca/Na}$ current caused by adrenaline during a 10-mV depolarising voltage-clamp pulse and the increase in the i_K -current tail recorded on return to the holding potential of -45 mV. These results correlate with previous findings that adrenaline increases both the slow inward and the plateau outward current in other cardiac tissues (for review see ref. 11).

As indicated in Fig. 1, the positive chronotropic effect of adrenaline is mediated by an increased rate of pacemaker depolarisation (PD). An increase of i_K will, however, tend to decelerate the PD. Although the larger $i_{Ca/Na}$ would by itself accelerate the PD in its later stages immediately preceding the upstroke, is it unlikely that combined with the augmented i_K it could account for the increased rate of change of voltage during the whole course of the PD seen in adrenaline.

What, then, is the mechanism of the increased pacemaker rate seen in adrenaline? We have found that an additional current, which has been previously noted in the rabbit SA node⁹ and in the frog sinus venosus¹² and which we will call i_t , is important both in normal pacemaker activity and in adrenaline-induced acceleration. Hyperpolarising voltage-clamp pulses from a holding potential of -42 mV (Fig. 3) reveal that i_t is a slowly increasing inward (or decreasing outward) current change whose amplitude increases with the clamp pulse amplitude. In Fig. 3 the voltage protocol has been superimposed on a record of

spontaneous activity in the same preparation to show the relevance of i_t to the potential range of the pacemaker depolarisation. Addition of adrenaline, 10^{-7} M, causes a substantial and reversible increase in the time-dependent change of i_t , as shown in Fig. 4. In the unclamped preparation such an increase would cause a faster PD, as observed during adrenaline action in Fig. 1.

Investigating the ionic nature of the i_t current system is difficult as it is not possible to hyperpolarise the small preparations used by more than 50 mV without risking membrane injury and consequent breakdown of the voltage clamp. We have therefore been unable to determine whether i_t has a reversal potential close to the equilibrium potential for K^+ ions, that is, at about -100 mV, which would show whether or not it could be equated with the i_{K2} current of the Purkinje fibre. Nevertheless, the resemblance to i_{K2} is striking: i_{K2} is deactivated by hyperpolarisations into the same voltage range as i_t and the change in i_t produced by adrenaline strongly resembles that produced in i_{K2} by a voltage shift of its kinetics. Regardless of whether it becomes possible to test this resemblance further experimentally, it is already clear that i_t has an important role both in normal rhythmic activity in the SA node and in mediating the response to adrenaline.

The financial support of the MRC and of the Wellcome Trust is acknowledged.

H. F. BROWN
D. DiFRANCESCO
S. J. NOBLE

University Laboratory of Physiology,
Parks Road, Oxford, UK

Received 10 May; accepted 4 June 1979.

1. Hauswirth, O., Noble, D. & Tsien, R. W. *Science* **162**, 916-917 (1968).
2. Vassort, G. *et al.* *Pflügers Arch. ges. Physiol.* **309**, 70-81 (1969).
3. Brown, H. F. & Noble, S. J. *J. Physiol., Lond.* **238**, 51-53P (1974).
4. Reuter, H. & Scholz, H. *J. Physiol., Lond.* **264**, 49-62 (1977).
5. Kohlhardt, M., Figulla, H. R. & Tripathi, O. *Basic Res. Cardiol.* **71**, 17-25 (1976).
6. Brown, H. F., Giles, W. & Noble, S. J. in *The Sinus Node* (ed. Bonke, F. I. M.) (Martinus Nijhoff, The Hague, 1978).
7. Brown, H. F., DiFrancesco, D. & Noble, S. J. *J. Physiol., Lond.* **290**, 31P-32P (1979).
8. Noma, A., Yangihara, K. & Irisawa, I. *Pflügers Arch. ges. Physiol.* **372**, 43-51 (1978).
9. Noma, A. & Irisawa, I. *Pflügers Arch. ges. Physiol.* **364**, 45-52 (1976).
10. DiFrancesco, D., Noma, A. & Trautwein, W. *Pflügers Arch. ges. Physiol.* (in the press).
11. Brown, H. F., DiFrancesco, D. & Noble, S. J. in *Biological Oscillators* (eds Berridge, M. J. & Rapp, P. E.) (Cambridge University Press, in the press).
12. Brown, H. F., Giles, W. & Noble, S. J. *J. Physiol., Lond.* **271**, 783-816 (1977).
13. Noma, A. & Irisawa, I. *Pflügers Arch. ges. Physiol.* **366**, 251-258 (1976).

Role of the dendritic release of dopamine in the reciprocal control of the two nigro-striatal dopaminergic pathways

USING cats implanted with push-pull cannulae, we have previously demonstrated a reciprocal regulation of the activity of the two nigro-striatal dopaminergic pathways. Asymmetric fluctuations in the spontaneous release of dopamine (DA) were simultaneously seen in the two caudate nuclei (CN)¹. As revealed by the changes in DA release from nerve terminals, the unilateral nigral application of DA¹ or of drugs enhancing the dendritic release of DA, such as amphetamine or benzotropine (2), reduced the activity of ipsilateral dopaminergic neurones and induced an opposite effect in the contralateral side. Conversely, the unilateral nigral blockade of dopaminergic transmission by local application of neuroleptics enhanced the release of DA in the ipsilateral CN and decreased the transmitter release in the contralateral structure². Opposite changes in the release of DA were also seen in the two CN under unilateral delivery of sensory stimuli³ or during unilateral electrical stimulation of the cerebellar dentate nucleus⁴. These latter effects were associated with asymmetric changes in the dendritic release of DA in the two substantia nigrae (SN) which were in an opposite direction to those seen in the two CN (an increased

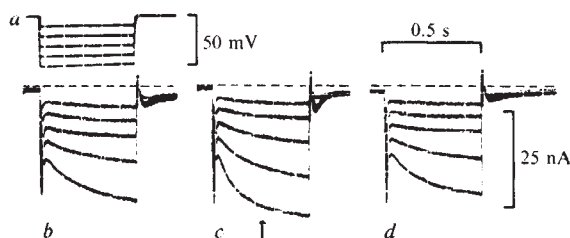


Fig. 4 Increase of i_t in adrenaline. i_t changes are activated by voltage-clamp hyperpolarisations from a holding potential of -36 mV. *a*, Voltage protocol used throughout the experiment; *b*, control; *c*, during adrenaline, 10^{-7} M; *d*, return.

release of DA from nerve terminals corresponded to a decreased release of the transmitter from dendrites and vice versa). These experiments indicated that DA which is released from dendrites in one SN regulates the activity of the ipsilateral dopaminergic neurones; they also suggested that it contributes to the control of the contralateral dopaminergic neurones by influencing the dendritic release of DA in the contralateral SN. This hypothesis was confirmed in the present study by measuring the changes in DA release from nerve terminals and dendrites of the two pathways under pharmacological blockade or facilitation of dopaminergic transmission in one SN.

Adult cats of both sexes were anaesthetised with halothane and implanted with four push-pull cannulae (0.9 mm diameter), one in each SN and each CN, to measure continuously the release of ^3H -DA synthesised from ^3H -tyrosine, as previously described⁵. An artificial cerebrospinal fluid (CSF) containing ^3H -tyrosine (50 Ci mmole⁻¹, 50 $\mu\text{Ci ml}^{-1}$) was introduced in each structure at a rate of 1 ml h⁻¹ and ^3H -DA was estimated in serial 10-min superfusate fractions. ^3H -DA was separated from ^3H -tyrosine and ^3H -labelled metabolites by ion-exchange chromatography and alumina adsorption. α -Methyl-*p*-tyrosine (α -MT, 10^{-4}M), an inhibitor of DA synthesis, or (+)amphetamine (10^{-6}M) were applied to the left SN 22 h after the onset of superfusion of the four structures with ^3H -tyrosine to block or facilitate local dopaminergic transmission, respectively.

The unilateral nigral application of α -MT for 60 min markedly stimulated the release of ^3H -DA in the ipsilateral CN

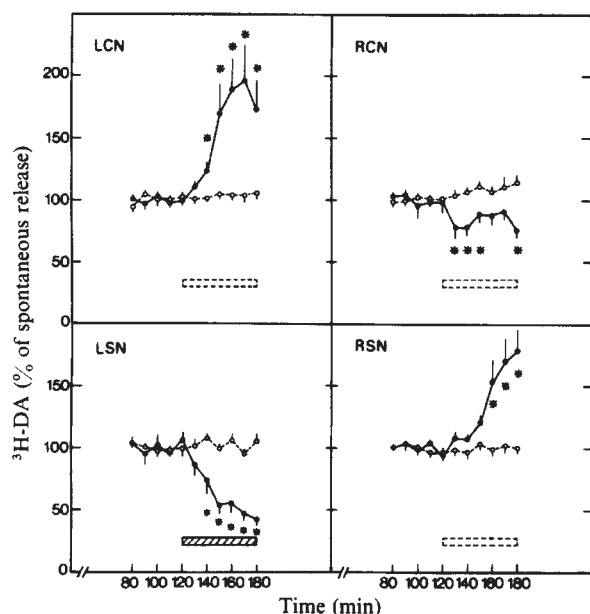


Fig. 1 Effects of α -methyl-*p*-tyrosine application into the left substantia nigra on the release of ^3H -dopamine from the two caudate nuclei and the two substantia nigrae. Four push-pull cannulae were simultaneously implanted in the left (LCN) and the right (RCN) caudate nuclei and in the left (LSN) and the right (RSN) substantia nigrae in anaesthetised cats. The four structures were perfused with an artificial CSF containing L[3,5- ^3H]tyrosine (50 Ci mmol⁻¹, 50 $\mu\text{Ci ml}^{-1}$, 1 ml h⁻¹). ^3H -Dopamine (^3H -DA) was estimated in 10-min successive superfusate fractions. The average quantities of ^3H -DA released in superfusate fractions of the SN (0.4 nCi) and the CN (0.6 nCi) were, respectively, 20 and 30 times the blank value. α -Methyl-*p*-tyrosine (α -MT, 10^{-4}M) was introduced for 60 min into the CSF superfusing the LSN (hatched bar). In each animal and for each cannula, ^3H -DA in each successive fraction was expressed as a percentage of an average spontaneous release calculated from the five fractions collected before the treatment. Data are the mean $\pm$ s.e.m. of results obtained with five animals (—). * $P < 0.05$ when compared with corresponding control values obtained in five untreated animals (---). ▨, α -MT application.

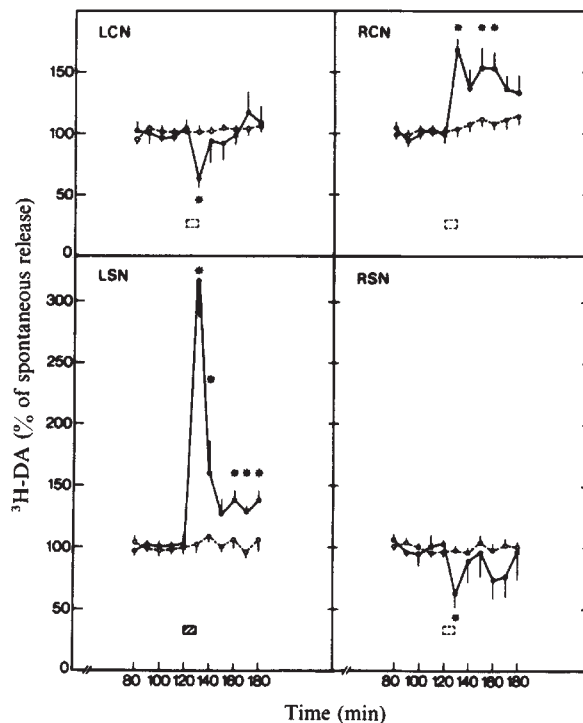


Fig. 2 Effects of (+)amphetamine application into the left substantia nigra on the release of ^3H -dopamine from the two caudate nuclei and the two substantia nigrae. Five animals were treated as described in Fig. 1 except that (+)amphetamine (10^{-6}M) was applied for 10 min into the left substantia nigra (LSN). Data are calculated and expressed as in Fig. 1. * $P < 0.05$ when compared with corresponding control values obtained in five untreated cats (---). ▨, (+)Amphetamine application.

and induced an opposite effect in the contralateral structure. Moreover, in contrast to that observed in the ipsilateral side, this treatment stimulated the dendritic release of ^3H -DA in the contralateral SN (Fig. 1). The unilateral nigral application of (+)amphetamine for 10 min not only induced asymmetric changes in the release of ^3H -DA in both CN (as previously shown²), but also in both SN. These effects were in opposite directions to those induced by α -MT (Fig. 2).

To confirm that the contralateral effects induced by the unilateral application of α -MT or (+)amphetamine were mediated by the changes in DA release induced in the ipsilateral SN and were not secondary to those observed in the ipsilateral CN, these drugs were applied to the left CN and ^3H -DA was simultaneously measured in both CN and both SN. α -MT (10^{-4}M) and (+)amphetamine (10^{-6}M) which, respectively, reduced and enhanced the local release of ^3H -DA in the left CN, failed to change ^3H -DA release in the contralateral CN or SN (Figs. 3, 4). Also, marked changes in the dendritic release of ^3H -DA were seen in the ipsilateral SN (Figs. 3, 4). They were parallel to those detected in the CN, as the dendritic release of ^3H -DA was reduced by α -MT and enhanced by (+)amphetamine.

Several conclusions on the influences of DA can be drawn from the present study and are summarised in Fig. 5. First, the activation of ipsilateral dopaminergic neurones induced by the unilateral application of α -MT demonstrates that DA released from dendrites tonically inhibits the activity of the ipsilateral neurones. This may mainly result from a reduced availability of DA at dopaminergic receptor sites located either on dopaminergic cells or their dendrites^{6,7} or on nigral afferent fibres^{8,9} (Fig. 5a). Second, the effects of α -MT and (+)amphetamine on the dendritic release of ^3H -DA imply that these drugs, respectively, stimulated or inhibited the firing rate of the ipsilateral dopaminergic neurones through their effects on

the dendritic release of DA. Thus, the changes in DA dendritic release induced by local modifications of the dopaminergic transmission in the ipsilateral CN further support the postulated¹⁰ existence of a striato-nigral neuronal loop controlling the activity of the dopaminergic neurones (Fig. 5b). Third, the asymmetric changes in DA release seen in both CN under the unilateral application of various stimuli¹⁻⁴ cannot be related to modifications of ³H-DA release in one of the CN (Fig. 5c). Indeed, the changes in ³H-DA release from nerve terminals in one CN induced by the local application of α -MT or (+)-amphetamine did not affect ³H-DA release in the contralateral CN (Figs 3, 4). We therefore postulate that the asymmetric changes in DA release observed in the two CN result from the change in the dendritic release of DA occurring in one SN (Fig. 5d or e). This is well illustrated by the experiments shown in Figs 1 and 2. DA released from dendrites in one SN must thus

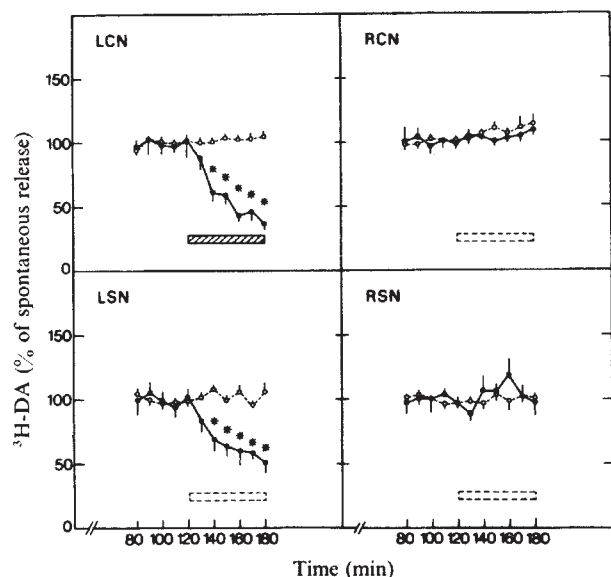


Fig. 3 Effects of α -methyl-*p*-tyrosine application into the left caudate nucleus on the release of ³H-dopamine from the two caudate nuclei and the two substantia nigrae. Five animals were treated as described in Fig. 1 except that α -methyl-*p*-tyrosine (α -MT, 10^{-4} M) was applied for 60 min into the left caudate nucleus (LCN). Data are calculated and expressed as in Fig. 1. * $P < 0.05$ when compared with corresponding control values obtained in five untreated cats (---). ▨, α -MT application.

regulate the activity of a nigral efferent pathway involved in the control of the activity of the contralateral dopaminergic neurones. This could occur as DA acts not only on the dopaminergic neurones in the pars compacta but also on other neurones in the pars reticulata¹¹. DA could also indirectly affect the activity of the pars reticulata neurones by modifying the release of transmitters from nigral afferent fibres¹². As cells of the pars reticulata project to several structures in the brain¹³⁻¹⁷, the pathway involved remains to be determined. However, we can still discuss whether the final input of this pathway occurs in the contralateral SN (Fig. 5e) or contralateral CN (Fig. 5d). Fourth, the effects of unilateral nigral α -MT or (+)-amphetamine indicate that the asymmetric changes in the dendritic release of DA in both SN are responsible for the opposite asymmetric changes in DA release in both CN. Indeed, the changes in the dendritic release observed in the right SN during the application of these drugs in the left SN cannot be related to the opposite changes in ³H-DA release from nerve terminals seen in the contralateral CN (Fig. 5d). If this was the case, parallel changes should occur in the CN and the SN as observed during the local application of (+)-amphetamine or α -MT in the CN (Figs 3, 4). As there is apparently no direct anatomical connection between

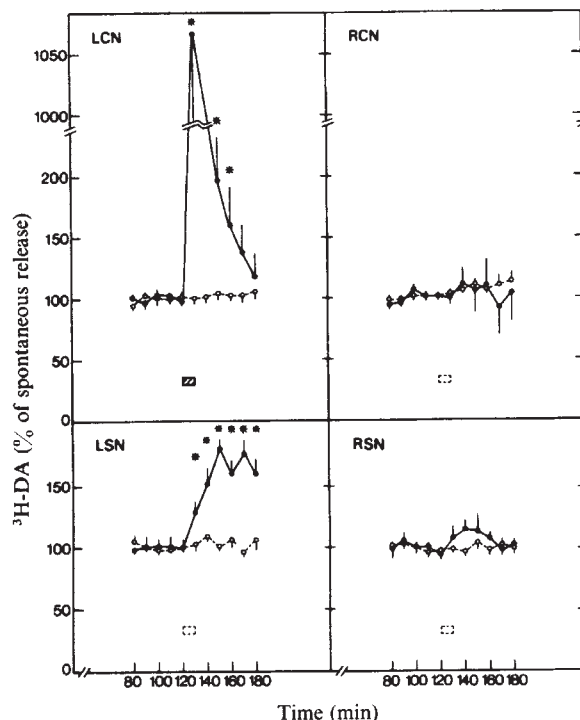


Fig. 4 Effects of (+)amphetamine application into the left caudate nucleus on the release of ³H-dopamine from the two caudate nuclei and the two substantia nigrae. Seven animals were treated as described in Fig. 1 except that (+)amphetamine (10^{-6} M) was applied for 10 min into the left caudate nucleus (LCN). Data are calculated and expressed as in Fig. 1. * $P < 0.05$ when compared with corresponding control values obtained in five untreated cats (---). ▨, (+)Amphetamine application.

the two SN, a polysynaptic pathway activated or inhibited by DA released from dendrites in one SN must be involved in the reciprocal control of the dendritic release of DA in the contralateral SN and thus in the regulation of the activity of the contralateral dopaminergic neurones (Fig. 5e).

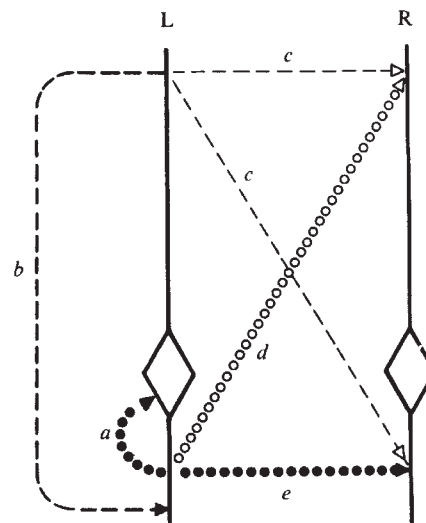


Fig. 5 Influences of dopamine released from dendrites and nerve terminals. Possible direct and indirect influences of dopamine (DA) released from nerve terminals (---, ---) or dendrites (○, ●) of the left (L) dopaminergic neurones on both nigro-striatal dopaminergic pathways (L, R). Our results indicate the existence of the influences *a*, *b* and *e*, and exclude *c* and *d* (see text).

Finally, we should consider why no contralateral effects were observed during the local application of α -MT or (+)amphetamine in the left CN despite the changes in the dendritic release of ^3H -DA induced in the ipsilateral SN. Recent experiments suggest that the striato-nigral substance P neurones regulate the activity of the dopaminergic neurones by acting on DA dendritic release¹⁸. Moreover, in these experiments, there was no effect in the contralateral SN and CN¹⁹, as was the case when dopaminergic drugs were applied in the left CN (Figs. 3, 4). Therefore, the striato-nigral substance P neurones may not only regulate the activity of the ipsilateral nigro-striatal dopaminergic neurones but could also influence the nigral cells involved in the reciprocal regulation of the two dopaminergic systems, and thus counteract the effect of DA on these cells. In fact, substance P nerve terminals are distributed in both the pars compacta and the pars reticulata²⁰⁻²² and substance P activates most cells in the SN^{23,24}. Furthermore, in contrast to that observed with substance P and its antibody, the unilateral nigral application of agonists or antagonist of γ -aminobutyric acid affected the release of DA in the contralateral CN²⁵. These effects were distinct from those induced by dopaminergic drugs, as in all cases the changes in ^3H -DA release seen in the two CN were symmetric.

Regarding their role in the coordination of sensory motor processes, it has previously been assumed that the nigro-striatal dopaminergic neurones act merely by modulating the activity of striatal cells. The demonstration of the dendritic release of DA, its important fluctuations in various physiological and pharmacological situations, and its role in regulating the activity, not only of the dopaminergic neurones but also of other nigral efferent pathways, indicate that the nigro-striatal dopaminergic neurones also control the delivery of messages originating from the SN. Such processes, which could be shared by other neurones, will undoubtedly influence our view of the transfer of information in the central nervous system.

This study was supported by grants from INSERM (ATP 71.78.103), DRET (78.004) and Rhône-Poulenc. A.C. is a research fellow of Rhône-Poulenc. We thank Miss F. Artaud for her excellent assistance.

V. LEVIEL
A. CHÉRAMY
J. GLOWINSKI

Groupe NB, INSERM U.114, Collège de France,
11 place Marcelin Berthelot,
75231 Paris, Cédex 5 France

Received 20 March; accepted 25 May 1979.

- Nieoullon, A., Chéramy, A. & Glowinski, J. *Science* **198**, 416-418 (1977).
- Nieoullon, A., Chéramy, A., Leviev, V. & Glowinski, J. *Eur. J. Pharmac.* **53**, 289-296 (1979).
- Nieoullon, A., Chéramy, A. & Glowinski, J. *Nature* **269**, 340-342 (1977).
- Nieoullon, A., Chéramy, A. & Glowinski, J. *Brain Res.* **148**, 143-152 (1978).
- Nieoullon, A., Chéramy, A. & Glowinski, J. *J. Neurochem.* **28**, 819-828 (1977).
- Bunney, B. S., Aghajanian, G. K. & Roth, R. H. *Nature new Biol.* **245**, 123-125 (1973).
- Groves, P. M., Wilson, C. J., Young, S. J. & Rebec, G. V. *Science* **190**, 522-529 (1975).
- Prémont, J. *et al. FEBS Lett.* **68**, 99-104 (1976).
- Gale, K., Guidotti, A. & Costa, E. *Science* **195**, 503-505 (1977).
- Carlsson, A. & Lindqvist, M. *Acta pharmac. tox.* **20**, 140-144 (1963).
- Dray, A., Gonye, T. J., Oakley, N. R. & Tanner, T. *Brain Res.* **113**, 45-57 (1976).
- Reubi, J. C., Iversen, L. L. & Jessell, T. M. *Nature* **268**, 652-654 (1977).
- Llamos, A. & Reinoso-Suarez, F. in *3rd Symp. on Parkinson Disease* (eds Gillingham, F. J. & Donaldson, I. L. M.) 82-87 (Livingstone, London, 1969).
- Rinvik, E. *Brain Res.* **91**, 313-318 (1975).
- Royce, G. J. *Brain Res.* **153**, 465-475 (1978).
- Sakai, K., Touret, M., Salvetti, D. & Jouve, M. in *Interaction between Putative Neurotransmitters in the Brain* (eds Garattini, S., Pujol, J. F. & Samanin, R.) 319-342 (Raven, New York, 1978).
- Chan-Palay, V. *Cerebellar Dentate Nucleus: Organization, Cytology and Transmitters*, 202 (Springer, Berlin, 1977).
- Chéramy, A. *et al. Brain Res.* **155**, 404-408 (1978).
- Michelot, R., Leviev, V., Giurgeff-Chesselet, M. F., Chéramy, A. & Glowinski, J. *J. Life Sci.* **24**, 715-724 (1979).
- Hökfelt, T., Kollerth, J. O., Nilsson, G. & Pernow, B. *Science* **190**, 889-890 (1975).
- Jessel, T. M., Emson, P. C., Paxinos, C. & Cuello, A. C. *Brain Res.* **152**, 487-498 (1978).
- Gauchy, C. *et al. Neurosci. Lett.* **12**, 127-131 (1979).
- Davies, J. & Dray, A. *Brain Res.* **107**, 623-627 (1976).
- Walker, R. J., Kemp, J. A., Yazima, H., Kitagawa, K. & Woodruff, G. N. *Experientia* **32**, 214-215 (1976).
- Leviev, V., Chéramy, A., Nieoullon, A. & Glowinski, J. *Brain Res.* (in the press).

The role of mitogenic lectins in T-cell triggering

THE process of T-lymphocyte activation requires the participation of metabolically active non- θ -bearing accessory cells¹. As first conclusively shown by Habu and Raff², this requirement is also true of lectin-dependent triggering. Thus, T-cell activation by mitogenic lectins depends on the presence of Ia-positive non-T cells^{3,4}. The mere binding of lectins such as concanavalin A (Con A) to the surface membrane of T cells does not trigger the cells to go through the mitotic cycle. This can only be achieved by the activity of growth factors present in conditioned media (CM) from Con A-stimulated cell cultures⁵. Such growth factors, however, cannot activate normal, resting T cells, although they are competent to maintain activated T cells in exponential growth for indefinite periods of time⁶. Hence, to be mitogenic a lectin must not only induce the *in situ* production of T-cell growth factors, but must also render resting T cells sensitive to the mitogenic activity of these growth factors. In this report we demonstrate that the binding of Con A to purified T cells can in a relatively short time, modify their functional sensitivity to growth factors, even though this binding is not sufficient to stimulate them to proliferate.

In our initial experiments we have used CM obtained at 24 h from Con A-stimulated spleen cell cultures. CM were always tested for T-cell growth factor activity on cultures of purified T-cell blasts⁴. Normal spleen cells, when cultured in such CM where the lectin has been inactivated (in our experiments by supplementation with the specific inhibitor sugar α -methyl-D-mannoside (α -MM)) are not induced to grow⁶ (Figs 1, 2 and Table 1). To demonstrate that direct interaction with mitogenic lectins renders T cells sensitive to growth factors, we incubated normal spleen cells in medium, either alone or with $5 \mu\text{g ml}^{-1}$ Con A, for various periods of time. The cells were then washed

Table 1 Direct interaction of Con A with T cells induces responsiveness to growth factors by expression of acceptor sites

Expts 1 and 2	Additions to cultures	Spleen cell responses (c.p.m. per culture $\times 10^{-3}$) after:			
		Anti-Ig column		Anti-I-A treatment and anti-Ig column	
		Expt 1	Expt 2	Expt 1	Expt 2
	None	1.6 $\pm$ 0.2	1.5 $\pm$ 0.2	0.3 $\pm$ 0.2	0.4 $\pm$ 0.3
	Con A	339.3 $\pm$ 11.2	165.2 $\pm$ 4.9	3.6 $\pm$ 0.4	2.6 $\pm$ 0.6
	CM	ND	3.7 $\pm$ 0.3	2.1 $\pm$ 0.4	0.8 $\pm$ 0.3
	Con A + CM	ND	81.4 $\pm$ 5.7	79.4 $\pm$ 3.9	45.9 $\pm$ 3.9
Expts 3 and 4		T-cell blast response (c.p.m. per culture $\times 10^{-3}$) to CM absorbed on:			
		Expt 3		Expt 4	
	None	—	1.8 $\pm$ 0.4	0.6 $\pm$ 0.4	
	Con A	—	1.3 $\pm$ 0.2	1.3 $\pm$ 0.4	
	CM	—	25.8 $\pm$ 2.7	18.6 $\pm$ 1.9	
	CM	Normal spleen cells	19.3 $\pm$ 2.5	15.4 $\pm$ 2.1	
	CM	4-h Con A spleen cells	12.9 $\pm$ 1.6	7.1 $\pm$ 0.9	
	CM	T-cell blasts	9.2 $\pm$ 2.0	0.7 $\pm$ 0.2	

In expts 1 and 2 normal C3H/HeJ spleen cells either untreated or incubated for 30 min in the cold with monoclonal anti-I-A (clone H-116/32, see ref. 14) antibodies were passed over anti-Ig columns (13) and then stimulated in culture ($2 \times 10^5 \text{ ml}^{-1}$) with Con A ($5 \mu\text{g ml}^{-1}$) or a standard preparation of CM at a 50% (v/v) final concentration. ^3H -Thymidine uptake was measured on day 3 of culture. This type of experiment has been reproduced on six independent occasions. In expts 3 and 4 T-cell blasts kept in standard CM for over 2 weeks were cultured ($5 \times 10^4 \text{ ml}^{-1}$) in the presence of either Con A ($5 \mu\text{g ml}^{-1}$) or CM (50%, v/v, final concentration) which had been absorbed for 30 min in the cold with the indicated cells ($2 \times 10^8 \text{ ml}^{-1}$ CM). Cells which had been exposed to Con A for 4 h were washed five times with 20 mg ml^{-1} α -MM before use in the absorption. T-cell blast responses were measured on day 2 after initiation of culture. Mixing experiments showing that the loss of activity in absorbed CM is not due to inhibitory activities have been carried out and partially presented elsewhere⁶. All results are shown as the mean values and s.d. from triplicate cultures. ND, not determined.

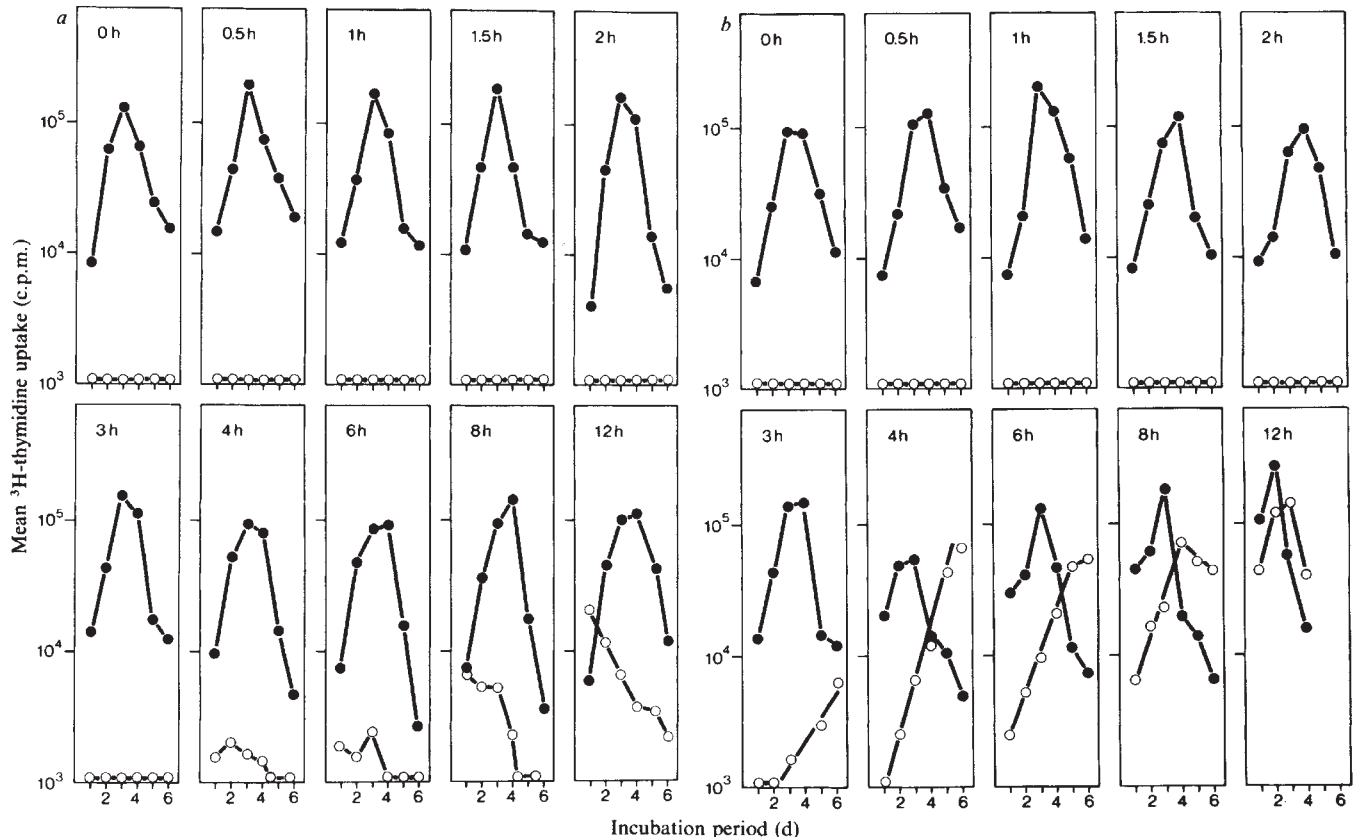


Fig. 1 Induction of responsiveness to growth factors by a short pulse of Con A. C3H/HeJ spleen cells ($5 \times 10^6 \text{ ml}^{-1}$) were incubated for the indicated periods of time, at 37°C , in culture medium [RPMI-1640 containing $50 \mu\text{g ml}^{-1}$ neomycin, 10 mM HEPES, $5 \times 10^{-5} \text{ M}$ 2-mercaptoethanol and 10% fetal calf serum (GIBCO, batch L3644015)] in the presence or absence of $5 \mu\text{g ml}^{-1}$ Con A (Pharmacia). The cells were then washed five times in medium containing 20 mg ml^{-1} α -MM and recultured ($2 \times 10^5 \text{ ml}^{-1}$) either in medium or in growth-factor-containing medium (a 50% (v/v) concentration of CM obtained from 24-h cultures of Con A-stimulated spleen cells, supplemented with 20 mg ml^{-1} α -MM) or in medium containing $5 \mu\text{g ml}^{-1}$ Con A. The proliferative responses in triplicate cultures were assayed every day as indicated after a 2-h pulse of ^3H -thymidine (Amersham TRK310; specific activity 2 Ci mmol^{-1} ; $5 \mu\text{Ci ml}^{-1}$). *a*, Preincubation in Con A, cultures in Con A (●) or medium (○). *b*, Preincubation in Con A, cultures in CM (○); all cultures with cells preincubated in medium and set up in either medium or CM had values below $1,000 \text{ c.p.m.}$ and are not shown. These cells were, however, functionally competent, as demonstrated by their responsiveness to Con A (●) in *b*. The results represent the mean values observed in triplicate cultures; the bars indicating s.d. are smaller than the symbols in the figure. Identical results were observed in three independent experiments.

free of lectin and their reactivity to growth factors was assessed in secondary cultures.

As shown in Fig. 1, incubation of spleen cells with optimal mitogenic doses of Con A for less than 8 h does not result in initiation of growth; this is a good indication that the removal of lectin after preincubation was effective. Exposure to Con A for 8 or 12 h results in a limited degree of cell activation to DNA synthesis, which, however, steadily decreases from day 1. In contrast, Con A pulses as short as 3 h result in the acquisition of responsiveness to growth factors, as demonstrated by the exponential growth observed in such cultures from days 1 to 6. Note that the response to Con A in the secondary cultures is the same whether the cells had been preincubated in the presence or absence of Con A. These experiments demonstrate that normal splenic T cells, which do not respond to either CM or a short pulse of Con A, initiate exponential growth when cultured in CM after the Con A pulse.

As pointed out above, T-cell proliferation depends on growth factors which are induced by mitogenic lectins⁴. We have analysed in detail the kinetics of growth factor production to separate the two triggering events: induction of sensitivity to growth factors (shown in Fig. 1), and production of these growth factors. Normal spleen cells were cultured with $5 \mu\text{g ml}^{-1}$ Con A, and the culture supernatants collected after various incubation periods and tested for their content in T-cell blast specific growth-factor activity, as described elsewhere⁴. As shown in Fig. 3, before 6 h no growth-promoting activity was detected in the supernatants, and this activity became maximal only between 12 and 24 h.

The difference in the kinetics of these two fundamental steps in T-cell triggering should be remembered when interpreting published observations on the 'commitment' of T lymphocytes to mitosis after mitogen pulses⁷⁻¹². It is clear that in any experimental conditions the limiting factor is the available concentration of growth factors; therefore, such experiments provide information on the conditions for the generation of these molecules.

As we postulate two independent events in the pathway of lectin-dependent T-cell triggering, and as the whole activation process is dependent on Ia-positive accessory cells²⁻⁴, we investigated whether or not both steps show similar cellular requirements. Normal spleen cells were passed over Ig-anti-Ig columns¹³ with or without previous treatment with monoclonal mouse anti-I-A antibodies¹⁴. Effluent cells from the columns were then either simply cultured with Con A (to assess the effectiveness of accessory cell depletion) or tested for their ability to acquire responsiveness to growth factors after a short exposure to Con A. As we reported previously⁴ and as shown in Table 1, although depletion of B cells has no influence on the T-cell response to Con A, concomitant depletion of I-A-bearing cells results in the abolition of the responses. Such responses are, however, largely restored when preformed growth factors are added to the cultures. As already pointed out, the factors are not mitogenic in the absence of Con A (expts. 1, 2). These results indicate that Con A responses are dependent on the presence of accessory cells, because these Ia-positive cells are necessary for the generation of growth factors.

The experiment presented in Fig. 2 shows that the first step in

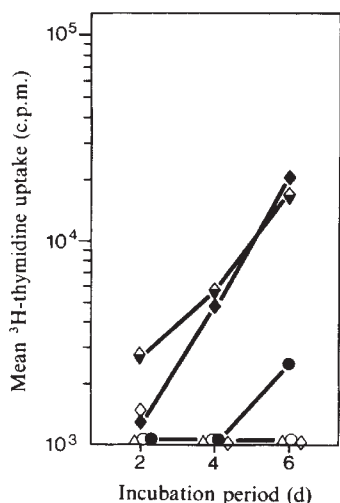


Fig. 2 Direct induction of responsiveness to growth factors in purified T cells by a short pulse of Con A. C₃H/HeJ spleen cells were treated with anti-I-A antibodies and passed through an anti-Ig column as described in Table 1, then cultured ($2 \times 10^5 \text{ ml}^{-1}$) in normal medium (○) or in medium containing $5 \mu\text{g ml}^{-1}$ Con A (●), a 50% final concentration of CM (◇) or both Con A and CM (◆). The same cells were also exposed to $5 \mu\text{g ml}^{-1}$ Con A for 4 h, washed five times in medium containing 20 mg ml^{-1} α -MM and cultured in medium (△) or in CM (◆).

triggering is independent of accessory cells. Purified T lymphocytes, depleted of Ia-positive cells, which do not respond to growth factors, acquire responsiveness to such growth factors during a 4 h treatment with Con A. After this short pulse, Con A is no longer required, but cell growth depends entirely on the presence of growth factors. We conclude that a direct interaction of the mitogenic lectin with the responding T cells is initially necessary for induction. Furthermore, this interaction is independent of the requirement for lectin and accessory cells in the generation of growth factors.

As we have shown before, responsiveness to growth factors is the result of the expression on the surface membrane of an acceptor site for those molecules⁶. We investigated whether or not cells pulsed with Con A for 4 h (which no longer required

Con A for responding to growth factors) already bear such acceptor sites on the surface membrane. Acceptor sites were assayed by their ability to absorb out T-cell-specific growth-promoting activity from supernatants. Con A-pulsed cells were compared with normal spleen cells and with T-cell blasts. As shown in Table 1 (expts 3, 4), spleen cells exposed to Con A for 4 h have more acceptor sites for growth factors than normal spleen cells, but not as many as do full blast cells, which are, of course, much larger.

The present results demonstrate that direct interaction of a mitogenic lectin with purified T-cell populations does not trigger them into proliferation, but it induces, in a relatively short time, surface membrane expression of functional receptors for growth factors. In addition, the mitogenic lectin induces the production of growth factors in a process which depends on accessory, Ia-positive cells.

Assuming that T-cell triggering, in both allogenic and restricted responses, is not fundamentally different from the lectin-dependent responses studied here, we may speculate on the physiological significance of the two independent events which seem to be strictly necessary for induction. Thus, the bi-functional role of mitogenic lectins may be equivalent to the role of antigen and major histocompatibility complex products on accessory cells in the process of specific T-cell induction.

Mouse anti-I-A antibodies were provided by Dr H. Lemke.

EVA-LOTTA LARSSON

Department of Biochemistry,
Arrhenius Laboratory,
University of Stockholm,
S-106 91 Stockholm, Sweden
and
Basel Institute for Immunology

ANTONIO COUTINHO

Basel Institute for Immunology,
Postfach, CH-4005 Basel, Switzerland

Received 29 January; accepted 1 June 1979.

1. Lafferty, K. J. & Woolnough, J. *Immun. Rev.* **35**, 231 (1977).
2. Habu, S. & Raff, M. C. *Eur. J. Immun.* **7**, 451 (1977).
3. Ahmann, G. B., Sachs, D. H. & Hodes, R. J. *J. Immun.* **121**, 1981 (1978).
4. Andersson, J., Grönvik, K., Larsson, E.-L. & Coutinho, A. *Eur. J. Immun.* (in the press).
5. Gillis, S. & Smith, K. A. *Nature* **268**, 154 (1977).
6. Coutinho, A., Larsson, E.-L., Grönvik, K. & Andersson, J. *Eur. J. Immun.* (in the press).
7. Powell, A. E. & Leon, M. A. *Expl Cell Res.* **62**, 315 (1970).
8. Novogrodsky, A. & Katcholski, E. *Biochim. biophys. Acta* **228**, 579 (1971).
9. Gunther, G. R., Wang, J. L. & Edelman, G. M. *J. Cell Biol.* **62**, 366 (1974).
10. Toyoshiwa, S., Iwata, M. & Osawa, T. *Nature* **264**, 447 (1976).
11. Milner, S. M. *Nature* **268**, 442 (1977).
12. Stenzel, K. H., Rubin, A. L. & Novogrodsky, A. *Expl Cell Res.* **115**, 285 (1978).
13. Wigzell, H., Sundquist, K. G. & Yoshida, T. O. *Scand. J. Immun.* **1**, 75 (1972).
14. Hämmerling, G. J. *et al. Curr. Topics Microbiol. Immun.* **81**, 100 (1978).

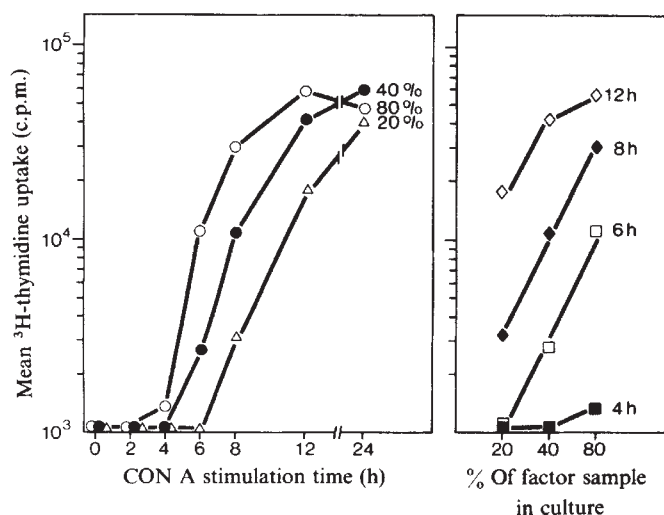


Fig. 3 Induction of T-cell growth factors in spleen cell cultures stimulated by Con A. C₃H/HeJ spleen cells ($5 \times 10^5 \text{ ml}^{-1}$) were cultured as described in Fig. 1 legend, in the presence of $5 \mu\text{g ml}^{-1}$ Con A. At the indicated times, CM from these cultures were collected, supplemented with 20 mg ml^{-1} α -MM and tested for growth-promoting activity in T-cell blasts which had been grown in standard CM for over 2 weeks. The test cultures contained 10^4 blasts in 0.2 ml and were supplemented by 80% (○), 40% (●) or 20% (△) concentrations of each supernatant. The results shown are the mean ^3H -thymidine uptake in triplicate cultures 3 days later. In the right hand panel, the same results are plotted to show the titration of the growth-factor activity in CM recovered after the various incubation periods. ◇, 12 h; ●, 8 h; □, 6 h; ■, 4 h.

Expression of V-region-like determinants on Ig-negative precursors in murine fetal liver and bone marrow

IDIOTYPES, as defined by Oudin¹, are antigenic markers associated with antibodies produced by a single individual. The demonstration of 'individual antigenic determinants' on human myeloma proteins² introduced the concept of idiotype as a clonal marker of Igs. However, as the operational definition of an idiotype is based on the specificity of anti-idiotypic antisera, the extent of cross-reactivity of a particular idiotype, and therefore the definition of idiotype, depends on the criteria set for the preparation of reagents. In fact, a number of 'public' or 'cross-reactive' idiotypes are currently described which are produced by all individuals of a strain or species or even by different species³. Yet, idiotypes are exclusively considered as markers of

V-gene products defining groups of related Ig molecules. The basic concept of Jerne's network theory⁴, however, postulates that the universes of idiotypes and of antigens are identical. As a consequence, the finding of idio-type-cross-reactive determinants on non-Ig molecules should be the rule. We have observed⁵ that a class of polyclonally distributed mitogen receptors on splenic B cells bear determinants cross-reactive with the public idio-type of the J558 myeloma protein of BALB/c origin. We have investigated the possibility that this finding might be related to the mechanism of generation and regulation of the antibody repertoire. Having no indications on the growth regulation of B-cell precursors, we started to investigate the antigenic relationship between Ig idiotypes and mitogen receptors on those cells. We report here the cross-reactivity between several germ-line idiotypes and non-Ig surface structures expressed on cells present in haematopoietic tissues. We hypothesise that growth receptors on precursor B cells have determinants cross-reactive with germ-line V-gene products, and that precursor cells are driven and selected by interactions between these receptors and molecules with anti-idio-type specificity also belonging to an incipient network transmitted in the germline.

Fetal livers were collected from BALB/c or (DBA/2×C57BL/6) F₁ embryos from day 13 to 15 of timed pregnancies. Fetal liver at that age does not contain detectable numbers of B cells as defined by surface Ig expression or reactivity to lipopolysaccharide (LPS)^{6,7}. Adult bone marrow and spleen were obtained from BALB/c and C57BL/6 mice. We have used indirect immunofluorescence to detect the membrane expression of IgM and IgD, LPS-receptors⁸ and determinants cross-reactive with a number of idiotypes of likely germline origin⁹,

Table 1 Presence of idio-type-cross-reacting structures on mouse fetal liver, adult bone marrow and adult spleen cells

Reagent	Fetal liver (day 13–15)	% Positive cells		Adult spleen
		Adult bone marrow		
Anti-Igs	<0.04			
Anti- μ	<0.04	3.4		44.3
		Total	μ^+	
Anti-J558	0.73	4.0	1.0	6.9
Anti-MOPC460*	0.82	3.6	1.2	6.6
Anti-W3129	0.89	4.1	0.8	4.9
Anti-TEPC15	1.2	4.1	1.2	4.9
Anti-MOPC104E	0.60	ND		5.5
BALB/c				
IgG1 aSRBC	<0.05	ND		ND
Sp-2				
(IgG2b aSRBC)	<0.05	ND		<0.05
Normal A/J IgG	<0.05	ND		<0.1
Anti-LPS-R	<0.04	0.8	0.7	7.5
				all μ^+ 80% δ^+

Anti-Igs and anti- μ antibodies were labelled with tetramethylrhodamine isothiocyanate. All anti-idiotypic antisera and anti-LPS-R antiserum⁷ were labelled with trinitrophenylsulphonic acid and their binding to cells detected by exposure to fluorescein-conjugated anti-dinitrophenyl antibodies. Binding of guinea pig anti-MOPC104E was also detected with fluorescent rabbit anti-guinea pig Ig antibodies. The two methods of indirect immunofluorescence gave comparable results. Anti- δ antiserum was labelled with arsonate and detected with rhodamine-labelled rabbit anti-arsonate antibodies. Living cells in suspension in medium containing 10% fetal calf serum and 20 mM Na₂Na₃ were exposed to one of the anti-idiotypic reagents for 30 min in ice, washed and exposed to the detecting reagent (see refs. 5, 12 for details). All cell samples strained with anti-idiotypic or anti-LPS-R were counterstained with anti- μ . Spleen cell samples were also counterstained with anti- δ . The stained samples were observed under a fluorescence microscope (Zeiss Photomikroskop II or Leitz Orthoplan), equipped with vertical illuminators and combinations of filters and dichroic mirrors specific for fluorescein and rhodamine. Total cells in each field were counted in phase contrast. At least 2,500 cells in fetal liver samples, and 1,000 cells in bone marrow and spleen samples were counted in each experiment and for each reagent. The figures reported here are average values of six experiments with fetal liver cells, three experiments with bone marrow, and seven experiments with spleen cells from different strains, as we did not observe differences in frequency of positive cells among all the strains tested.

* The three different anti-MOPC460 idio-type antisera gave comparable results. ND, Not determined.

Table 2 Simultaneous presence of more than one idio-type-cross-reacting structure on the membrane of fetal liver and spleen cells

	Reagents	% Positive cells	
		Fetal liver	Spleen
Anti-MOPC460 +anti-W3129	Total MOPC460	0.44	6.5
	'Pure' MOPC460	0.18	3.4
	Doubles	0.26	3.1
	'Pure' W3129	0.27	3.0
	Total W3129	0.53	6.1
Anti-J558 +anti-W3129	Total J558	0.62	6.6
	'Pure' J558	0.25	1.7
	Doubles	0.37	4.9
	'Pure' W3129	0.17	1.4
	Total W3129	0.54	6.3
Anti-TEPC15 +anti-W3129	Total TEPC15	0.77	ND
	'Pure' TEPC15	0.62	
	Doubles	0.15	
	'Pure' W3129	0.40	
	Total W3129	0.55	
Anti-MOPC460 +anti-MOPC104E	Total MOPC460	0.39	ND
	'Pure' MOPC460	0 (<0.05)	
	Doubles	0.39	
	'Pure' MOPC104E	0.25	
	Total MOPC104E	0.64	
Anti-J558 +anti-MOPC104E	Total J558	0.60	6.8
	'Pure' J558	0.10	5.5
	Doubles	0.50	1.3
	'Pure' MOPC104E	0.25	4.9
	Total MOPC104E	0.76	6.2
Anti-TEPC15 +anti-MOPC104E	Total TEPC15	0.73	ND
	'Pure' TEPC15	0.07	
	Doubles	0.64	
	'Pure' MOPC104E	0.06	
	Total MOPC104E	0.70	

Anti-W3129 was labelled with arsonate and detected with rhodamine-labelled rabbit anti-arsonate antibodies. Anti-MOPC104E was used unlabelled and detected with rhodamine-labelled rabbit anti-guinea pig antibodies. Anti-J558, anti-TEPC15 and anti-MOPC460 were labelled with trinitrophenylsulphonic acid and detected with fluorescein-labelled anti-dinitrophenyl. The cell suspensions were exposed to a pair of anti-idiotypic antibodies for 30 min in the cold, washed and exposed to a mixture of the fluorescent detecting reagents. Positive cells out of the total cell population were counted as indicated in Table 1 legend. For the evaluation of 'double' versus 'pure' cells, each field was observed under specific illumination for fluorescein and for rhodamine. At least 100, in most cases 200, positive cells were scored in each combination.

namely the idiotypes on the following BALB/c myeloma proteins: J558 ($\alpha\lambda_1$, anti- α -1,3 dextran), W3129 ($\alpha\kappa$, anti- α -1,16 dextran) TEPC15 ($\alpha\kappa$, anti-phosphorylcholine), MOPC460 ($\alpha\kappa$, anti-dinitrophenyl) and MOPC104E ($\mu\lambda_1$, anti- α -1,3 dextran). Antisera to the first three proteins were raised in mice^{5,10} while anti-MOPC104E was made in guinea pigs tolerant to mouse Ig. Three different anti-MOPC460 idiotypes were compared; conventional BALB/c, and rabbit antisera, and a monoclonal antibody¹². Although all antisera reacted only with the immunising and with idiotypically cross-reactive proteins, they were additionally absorbed on Sepharose-coupled IgM, IgA, IgG, κ and λ myeloma proteins and pooled mouse Igs. Anti-J558 and anti-MOPC460 idiotypes were antibodies purified by affinity chromatography. As controls, an IgG1 fraction of a BALB/c anti-SRBC antiserum, the monoclonal antibody Sp-2 (IgG2b anti-SRBC)¹³ and normal serum IgG from A/J mice were used. All reagents were labelled with trinitrophenyl¹⁴ or arsonate¹⁵ and used in indirect immunofluorescence^{5,16}.

Table 1 shows results of staining fetal liver, adult bone marrow and spleen cells with these reagents. All anti-idiotypes, but not normal IgG or anti-SRBC antibodies, stained a sizeable proportion of IgM-positive spleen lymphocytes (11–16% of all B cells). Nearly all cells stained by anti-idiotypic were $\mu^+\delta^-$, while the majority of cells expressing LPS-receptors were $\mu^+\delta^+$ suggesting that the former are less differentiated than the latter. A relatively small proportion of fetal liver cells (0.6–1.2%) were also stained by the anti-idiotypic reagents, while no cells expressed either Igs or LPS-receptors, or nonspecifically bound normal IgG. In bone marrow, anti-idiotypes stained ~4% of the total

cells, the majority of which were μ . None of the cells binding anti-idiotypes contained intracytoplasmic IgM, as demonstrated in experiments (not shown) where cells were stained on the membrane with anti-idiotypes, and in the cytoplasm, after fixation with anti- μ antibodies. It seems, therefore, that these cells are not those described in bone marrow and fetal liver by Raff *et al.*¹⁷.

Confidence on the specificity of the anti-idiotypic staining of Ig-negative fetal liver and bone marrow cells was provided by the lack of binding in the same staining conditions of other reagents such as mouse anti-sheep red blood cell (SRBC) or rabbit anti-LPS-receptor (LPS-R) antibodies. It is also clear that the molecules detected by anti-idiotypes on fetal liver are not idiotypes of passively acquired material IgG, as no cells were detectable in the same samples with anti-mouse Ig+L antibodies.

Further evidence for specificity of anti-idiotypic binding to target cells was provided by double-staining experiments shown in Table 2. Note that, although we consistently observed the presence of two such idiotypic-cross-reactive structures on the same cells both in spleen and fetal liver, a proportion of the positive cells was stained in all cases by only one reagent in the pair.

Table 3 shows results of experiments designed to characterise the membrane structures detected by some anti-idiotypes. The binding of anti-J558 and anti-MOPC104E idiotypes, but not that of anti-MOPC460 idiotypic, was specifically inhibited by α -1,3-containing dextran. These results suggest that the staining of fetal liver cells, as well as that of spleen cells⁵, by anti-J558 and anti-MOPC104E idiotypes is a specific reaction between anti-idiotypes and a membrane structure which is (or is associated with) a dextran binding site.

Having stressed the specificity of binding of purified anti-idiotypic antibodies to Ig-negative cells in fetal liver and bone marrow, we must discuss the specificity of these anti-idiotypic reagents. The antisera or purified antibodies used in these experiments are indeed anti-idiotypic when studied serologically for reactions with Ig molecules. On the other hand, the same reagents used to stain intracytoplasmic Igs of mature plasma cells in spleen, revealed frequencies of positive cells around 1–3%. These figures, although one order of magnitude lower than those observed by membrane staining of the same samples, were definitely higher than expected for 'individual specificities' and more compatible with V-region isotype specificities⁹. In fact, all the myeloma proteins used in these experiments carry 'public' idiotypes. It is likely therefore that the molecules we detect on the cell surface with these anti-idiotypes, do not cross-react with idiotypes in Oudin's *sensu stricto*, but rather with V-region determinants shared by related antibody molecules. Such V-region-like determinants are certainly not immunoglobulin in nature, or at least they are not associated with constant regions of known Ig isotypes (μ and L). In addition, the cross-reactivity observed here does not imply that we are dealing with products of Ig V-genes.

Functional studies in progress indicate that the cells described here are in fact B-cell precursors. Furthermore, the addition of

anti-idiotypes to fetal liver and bone marrow cultures can be shown to influence B-cell development, suggesting that such idiotypic-cross-reactive structures function as growth receptors. More than one specificity of idiotypic-cross-reactive structure can be found on the same cells, both in fetal liver and spleen, and this parallels previous findings of multiple mitogen receptor specificities on the same B cell¹⁸. This is in contrast with the strict monospecificity of Ig molecules of a single cell, where all membrane Igs, both IgM and IgD, bear the same idiotypic^{19,20}.

It is perhaps not fortuitous that such public idiotypic-cross-reactive structures are present on fetal liver and adult bone marrow Ig-negative B-cell precursors. If the repertoire of antibodies is a closed network⁴, and the germ-line specificities represent the main features of the system and are therefore anti-idiotypic, these 'idiotypes' on growth receptors of precursor cells would provide the internal conditions for inducing their proliferation, as well as a mechanism for selecting the desired idiotypic specificities. These are minimal requirements for somatic generation and regulation of antibody diversity, and they assume that the repertoire selection is guided by idiotypes and not by combining site specificities. This proposal assumes that physiological mitogens which drive precursor B-cell development are anti-idiotypes and not the ligands to which the idiotypic-bearing antibodies can bind. Only when the ligand can inhibit the reaction of anti-idiotypes with the triggering structures could we expect it to be mitogenic for precursors, as is the case for dextran^{18,21}. Speculations on the selective pressures which result in antibacterial specificity of growth receptors on B cells, as well as of most myeloma proteins with antibody activity, will be developed elsewhere.

We thank Dr G. Koehler for monoclonal antibody Sp-2, Dr S. Cammisuli for anti-arsenate antibodies, and Dr J. Haimovich for anti-dinitrophenyl antibodies. We are indebted to Dr Bonnie Blomberg for anti-J558 reagents and for helpful discussions.

LUCIANA FORNI

Basel Institute for Immunology

PIERRE-ANDRE CAZENAVE

Institut Pasteur,
Paris, France

HUMBERTO COSENZA

Depto de Microbiologia,
Universidad Nacional Autonoma de Honduras,
Tegucigalpa, Honduras

KLAUS FORSBECK

Department of Immunology,
Biomedicum, Uppsala, Sweden

ANTONIO COUTINHO

Basel Institute for Immunology,
Basel, Switzerland

Received 15 February; accepted 4 June 1979.

Table 3 Inhibition of the staining of fetal liver cells by anti-idiotypic antisera

Reagent	Inhibitor	% Positive cells	% Inhibition
Anti-J558	None	0.37	—
	Dextran B1355	0.04	89.2
	Dextran B512	0.36	2.8
Anti-MOPC104E	None	0.50	—
	Dextran B1355	0.18	64.0
	Dextran B512	0.37	26.0
Anti-MOPC460	None	0.46	—
	Dextran B1355	0.39	15.3
	Dextran B512	0.41	10.9

The staining was performed in the presence of dextran B1355 (containing α -1,3 and α -1,6 linkages) or dextran B512 (containing primarily α -1,6 linkages) at a concentration of 250 μ g ml⁻¹.

- Oudin, J. *Proc. R. Soc. Ser. B* **166**, 207 (1966).
- Slater, R. J., Ward, S. M. & Kunkel, H. G. *J. exp. Med.* **101**, 85 (1955).
- Ju, S. T., Benacerraf, B. & Dorf, M. E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6192 (1978).
- Jerne, N. K. *Ann. Immun., Inst. Pasteur* **125C**, 373 (1974).
- Coutinho, A., Forni, L. & Blomberg, B. *J. exp. Med.* **148**, 862 (1978).
- Edwards, G. E., Miller, R. G. & Phillips, R. A. *J. Immun.* **105**, 719 (1970).
- Melchers, F. *Eur. J. Immun.* **7**, 476 (1977).
- Forni, L. & Coutinho, A. *Eur. J. Immun.* **8**, 56 (1978).
- Potter, M. *Adv. Immun.* **25**, 141 (1977).
- Cosenza, H., Julius, M. H. & Augustin, A. A. *Immun. Rev.* **34**, 3 (1977).
- Eichmann, K. & Kindt, T. J. *J. exp. Med.* **134**, 532 (1971).
- Buttin, G. *et al. Curr. Topics Microbiol. Immun.* **81**, 27 (1978).
- Koehler, G. & Milstein, C. *Nature* **256**, 495 (1975).
- Little, J. R. & Eisen, H. N. in *Methods in Immunology and Immunochemistry* Vol. 1 (eds Williams, C. A. & Chase, M. W.) 131 (Academic, New York, 1967).
- Cammisuli, S. & Wofsy, L. *J. Immun.* **117**, 1965 (1976).
- Forni, L. & Coutinho, A. *Nature* **273**, 304 (1978).
- Raff, M. C., Megson, M., Owen, J. J. T. & Cooper, M. D. *Nature* **259**, 224 (1976).
- Gronowicz, E. & Coutinho, A. *Transplant. Rev.* **24**, 3 (1975).
- Salsano, F., Frøland, S. S., Natvig, J. B. & Michaelsen, T. E. *Scand. J. Immun.* **3**, 841 (1974).
- Fu, S. M., Winchester, R. J. & Kunkel, H. G. *J. Immun.* **114**, 250 (1975).
- Diamantstein, T., Blittstein-Willinger, E. & Schulz, G. *Nature* **250**, 596 (1974).

Tumour promoter TPA enhances transformation of human leukocytes by Epstein-Barr virus

TUMOUR-promoting phorbol esters have contributed to the concept of multiple-step carcinogenesis¹, as illustrated in the two-stage mouse skin carcinogenesis model². Promoters are not significantly carcinogenic when tested alone, but decrease the threshold dose of carcinogens (initiators) and reduce the latency period for tumour appearance. 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA) and several other diterpene esters represent potent tumour promoters³⁻⁵. Recently, it has been shown that these promoters efficiently induce persisting genomes of oncogenic herpesviruses, particularly of Epstein-Barr virus (EBV)⁶⁻⁸, resulting in high percentages of viral antigen and particle-positive cells. Other observations showed that TPA enhances *in vitro* transformation of fibroblasts previously exposed to polycyclic aromatic hydrocarbons, UV light or adenoviruses⁹⁻¹². These studies prompted us to analyse the effect of TPA not only on virus induction in EBV genome-harboring cells, but also on EBV-induced transformation of human cord blood lymphocytes. Our data show that low concentrations of TPA significantly enhance transformation of lymphocytes by EBV.

Initial experiments were designed to investigate the effect of various concentrations of TPA on colony formation of human cord blood leukocytes after infection by transforming EBV derived from the B95-8 line¹³. As previously described¹⁴⁻¹⁶, typical rod-shaped EBV-transformed colonies appeared after about 2 weeks of incubation. Figure 1 shows that cultures treated with quite a wide range of TPA concentrations showed a marked increase in the number of colonies when compared with the untreated controls. The optimal concentration of TPA was 0.31 ng ml⁻¹. In these conditions six times more colonies were counted than in cultures without TPA; transformed colonies appeared earlier and were larger in cultures kept in the presence

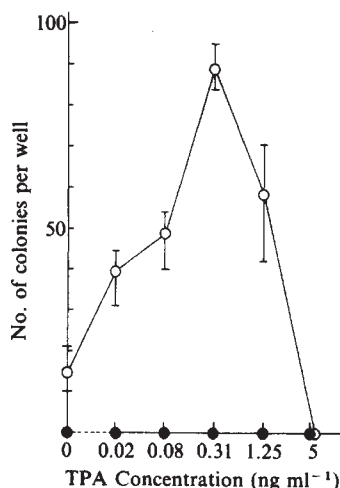


Fig. 1 Effect of TPA on colony formation of human cord blood lymphocytes after infection with EBV. Umbilical cord blood lymphocytes were prepared by a conventional Ficoll-Hypaque procedure²⁰, and infected with EBV derived from B95-8 cells. The virus concentration procedure has been described previously¹⁴. Colony formation by EBV-transformed cells in soft agar was determined as reported elsewhere¹⁵. The number of transformed cell colonies growing in agarose was quantitated under a low power microscope 2 weeks after infection and expressed as the mean value of four wells. The bars reveal the observed deviations. TPA was used in fourfold dilutions starting with 5 ng ml⁻¹. ●, Preparations without EBV; ○, preparations infected with EBV.

of TPA, although the morphology of TPA-treated and untreated cultures was similar.

TPA has been reported to stimulate T lymphocytes specifically¹⁷. However, no transformed lymphoblastoid colonies were observed in cultures treated exclusively with TPA, although some differentiated cell colonies were detected. As the colonies in TPA-treated cells were also EB nuclear antigen-positive, these results show that they were indeed transformed by EBV. Similar experiments were carried out using peripheral leukocytes of an EBV-seroreactive healthy adult. The results also confirmed that at an optimal concentration of 0.31 ng ml⁻¹ transformation efficiency was enhanced approximately fivefold (data not shown). Therefore, 0.31 ng ml⁻¹ was used in further experiments by modifying the concentration of the virus used for transformation tests. As shown in Fig. 2, the maximal number of colonies was obtained at a virus dilution of 1:4 in all preparations tested. At this virus concentration fivefold higher counts of colonies were noted in TPA-treated cultures. The number of colonies decreased proportionally to the virus dilution, starting at a dilution of 1:4 in both types of culture. The virus-dose response was similar to that observed in previous studies¹⁴. At higher virus concentrations the decrease in the number of transformed colonies may be attributed to cell killing activity of the concentrated preparations¹⁴.

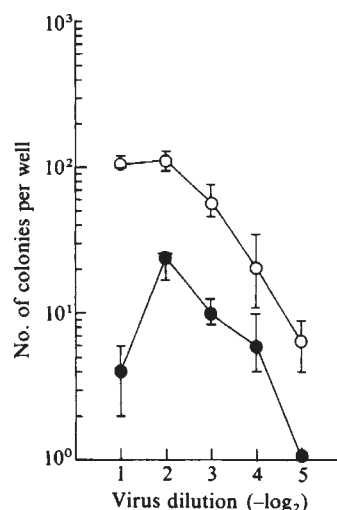


Fig. 2 Colony formation of human cord blood lymphocytes after infection with serial dilutions of EBV in the presence or absence of TPA. The procedures are the same as those outlined in Fig. 1. Cells are infected by serial dilutions of EBV and mixed with the agarose medium with (○) or without (●) 0.31 ng ml⁻¹ TPA.

We then determined whether TPA also stimulates colony formation of established lymphoblastoid lines in agarose even without the addition of exogenous EBV and whether this effect is specific for EBV genome-harboring cells. For this we used one EBV genome-negative line, BJAB, and four EBV-positive lines, P3HR-1, two Raji clones¹⁸ and a newly established human cord leukocyte line, CBL-1 (about 2 months after establishment). The data show that except for CBL-1, TPA did not enhance colony formation of these cells. There was even a reduction in the number of colonies (15–40% less than in the control) which were also more dispersed in the presence of TPA. However, TPA caused an approximately threefold increase in colony formation in CBL-1 cells (Fig. 3).

We also investigated the effect of TPA on transformation by end-point dilution. Again, TPA increased the transformation efficiency in liquid culture systems. In the presence of 0.31 ng ml⁻¹ TPA, two independent experiments yielded about fivefold higher TD₅₀ per ml (7.9×10^{-5} and 2.5×10^{-5}) than in the untreated controls (1.6×10^{-5} and 5.0×10^{-6}).

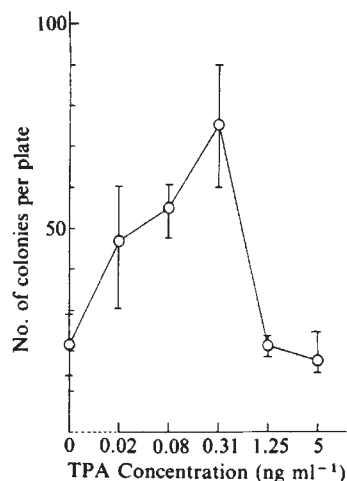


Fig. 3 Effect of TPA on cells of a newly established B95-8 EBV transformed cord blood lymphocyte line, CBL-1. The procedure for analysing the colony formation of cells from established lines was basically identical to that of the previous experiments¹⁸. 5,000 cells were seeded into each plate. The number of colonies was counted 7–9 days after inoculation. All values are the mean of three plates. Bars represent the deviations of these plates.

Our results indicate: (1) TPA at low molarity caused a three- to seven-fold enhancement of transformation of human leukocytes by EBV. (2) TPA does not increase the plating efficiency of three 'old' established cell lines in agarose regardless of their EBV genome state. (3) TPA enhances the plating efficiency of a newly established line in agarose. These results suggest that TPA acts on transformation steps rather than inducing nonspecific colony-stimulating factors that are known to promote cell growth in agarose¹⁹.

This work was supported by the Deutsche Forschungsgemeinschaft (Ha 449/12). N.Y. is a fellow of the Alexander von Humboldt-Stiftung.

NAOKI YAMAMOTO

HARALD ZUR HAUSEN

Institut für Virologie, Zentrum für Hygiene,
Universität Freiburg,
Hermann-Herder-Str. 11,
78 Freiburg, FRG

Received 26 April; accepted 13 June 1979.

- Foulds, L. *Neoplastic Development* (Academic, New York, 1969).
- Berenblum, I. in *Cancer Vol. 1* (ed. Becker, F. F.) 323–344 (Plenum, New York, 1975).
- Van Duuren, B. L. *Prog. exp. Tumor Res.* **11**, 31–68 (1969).
- Boutwell, R. K. *CRC Crit. Rev. Tox.* **2**, 419–443 (1974).
- Hecker, E. in *Handbuch der Allgemeinen Pathologie*, Vol. 4 (ed. Grundmann, E.) 651–675 (Springer, Berlin, 1975).
- zur Hausen, H., O'Neill, F. J., Freese, U. K. & Hecker, E. *Nature* **272**, 373–375 (1978).
- zur Hausen, H., Bornkamm, G. W., Schmidt, R. & Hecker, E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 782–785 (1979).
- Yamamoto, N., Bister, K. & zur Hausen, H. *Nature* **278**, 553–554 (1979).
- Lasne, C., Gentil, A. & Chouroulinkov, I. *Nature* **247**, 490–491 (1974).
- Mondal, S., Brankow, D. W. & Heidelberger, C. *Cancer Res.* **36**, 2254–2260 (1976).
- Mondal, S. & Heidelberger, C. *Nature* **260**, 710–711 (1976).
- Fisher, P. B., Weinstein, B., Eisenberg, D. & Ginsberg, H. S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2311–2314 (1978).
- Miller, G. & Lipmann, M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 190–194 (1973).
- Yamamoto, N. & Hinuma, Y. *Int. J. Cancer* **17**, 191–196 (1976).
- Katsuki, T., Hinuma, Y., Yamamoto, N., Abo, T. & Kumagai, K. *Virology* **83**, 287–294 (1977).
- Yamamoto, N. & Hinuma, Y. *Virology* **89**, 623–626 (1978).
- Touraine, J. L. *et al. J. exp. Med.* **145**, 460–465 (1977).
- Bister, K., Yamamoto, N. & zur Hausen, H. *Int. J. Cancer* **23**, 818–825 (1979).
- Macpherson, I. in *Tissue Cultures* (eds Kruse, P. F. & Patterson, M. K.) Ch. 7 (Academic, New York, 1973).
- Böyum, A. *Scand. J. chim. Invest.* **21**, Suppl. 97, 77–89 (1968).

Oxidant damage mediates variant red cell resistance to malaria

IT is thought that the deleterious properties of thalassaemia (thal) and glucose-6-phosphate dehydrogenase deficiency (G6PD⁻) are balanced in the population by a conferred protection against malaria, and that some characteristic of the variant red cell acts directly to inhibit the invasion or intracellular development of the malaria parasite. Indeed, such is the case for sickle cell disorders^{1,2}. Recently, Pasvol *et al.*³, having shown that fetal haemoglobin (HbF) (in both fetal and adult cells) inhibits malaria parasite development, have suggested that the slight persistence of HbF in children with thal trait is responsible for the selective advantage of the thal gene. They have reported no difference in parasite development in the red cells of adult carriers of thal. Evidence is presented here that both α - and β -thal trait red cells from adults are refractory to parasite development because of oxidant sensitivity, and, furthermore, that fetal red cells and G6PD⁻ cells are refractory for the same reason.

Decreased protection against oxidant damage has frequently been suggested as a mechanism by which G6PD⁻ cells could inhibit the optimal multiplication of the malaria parasite. Etkin and Eaton have demonstrated the production of H₂O₂ by a rodent malaria parasite⁴, and the insufficient protective capacity of the G6PD⁻ cell is well known⁵. The possibility that thalassaemic cells also inhibit parasite growth by their sensitivity to oxidant stress⁶ has not previously been suggested. To study the role of oxidant sensitivity in resistance to malaria, *Plasmodium*

Table 1 Sensitivity to increased oxygen of *Plasmodium falciparum* in G6PD⁻ and α - and β -thal red cells

RBC	Expt	Parasites per 100 red cells		Relative multiplication 30%/17%
		30% O ₂	17% O ₂	
Normal	1	3.2	9.3	0.34
	2	3.6	8.1	0.44
	3	2.1	5.0	0.42
	4	1.9	4.3	0.44
				0.41 ± 0.05
G6PD ⁻	1	0.4	4.4	0.09
	2	0.1	3.1	0.03
	3	1.2	4.6	0.26
	4	3.9	18.3	0.21
				0.15 ± 0.11
β -thal	1	0.3	1.8	0.17
	2	3.6	10.0	0.36
	3	0.6	4.1	0.15
	4	0.2	2.5	0.09
				0.19 ± 0.12
α -thal	1	1.3	8.2	0.16
	2	0.5	2.7	0.19
				0.18

Suspensions of variant and normal red cells (5% v/v) were inoculated with less than 1/30 cell volume of cells infected with *P. falciparum* (FMG/FCR-3). Cultures were maintained by the method of Trager and Jensen⁷ as described previously¹. After 2 d growth at 17% O₂, single cultures were split and incubated in either 17% or 30% O₂ for 2 additional days. Normal parasites per 100 red cells were counted on day 4 and the relative multiplication at 30% O₂ was calculated. Differences in the counts on day 4 between experiments result from different initial infections. Parasites with abnormal morphology at 30% O₂ were of a sufficient number to indicate that intracellular development was inhibited rather than invasion blocked. G6PD⁻ red cells were obtained by screening⁸ samples collected from male Negro blood donors. β -Thal red cells were from the parents of children attending a thalassaemia clinic (mostly mothers). α -thal red cells were verified by globin chain synthesis (H. Chang, New York Hospital). G6PD⁻ experiments 3 and 4 were on the same sample. All others are from different individuals.

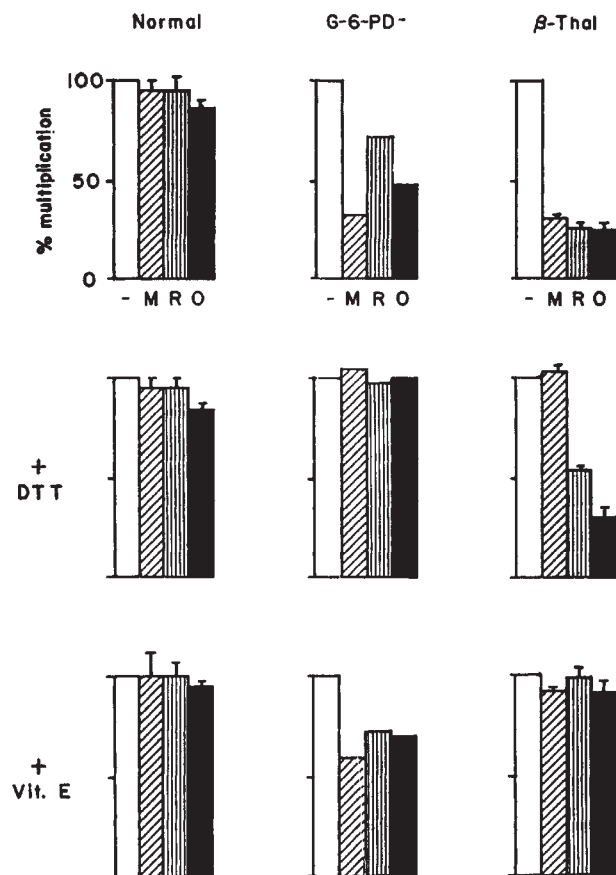


Fig. 1 Multiplication of *P. falciparum* in G6PD⁻ and β -thal red cells when oxidant damage is enhanced or inhibited. An experiment identical to that in Table 1 was carried out in the following conditions during the second 2 d: 17% O₂ (-), +2.5 μ M menadione (M), +2.0 μ M riboflavin (R), 25% O₂ (O). DTT (50 μ M) and vitamin E (0.2 mg ml⁻¹) were added as indicated. Height of the bars represents multiplication relative to that at 17% O₂. Error bars for normal and thal cells show one s.e.m. from three experiments (different donors). G6PD⁻ values are the means from two experiments. Neither DTT nor vitamin E had significant effects on growth in normal cells at 17% O₂.

falciparum, the human malaria parasite, was cultured in G6PD⁻, α - and β -thal trait, and HbF-containing erythrocytes in conditions which either enhanced or inhibited oxidant damage.

Cultures of G6PD⁻ and both α - and β -thal trait red cells in RPMI-1640 medium in an atmosphere of 17% O₂/3% CO₂ were inoculated with less than 1/30 volume of infected red cells. All variants supported normal multiplication of the parasite for 4 d in these conditions. The erythrocytic cycle of the parasite is 48 h and an eight- to tenfold multiplication was seen at this time. Thus, after 2 d, virtually all parasites had invaded variant cells. If cultures were then transferred to an atmosphere of 30% O₂/3% CO₂, an inhibition of multiplication occurred that was greater in G6PD⁻ and thal cells than in normal cells (Table 1) ($P < 0.01$). The response of these host-parasite systems to oxidative stress was further characterised by parallel incubation in either 25% O₂, 2 μ M riboflavin, or 2.5 μ M menadione (Fig. 1). Riboflavin was more inhibitory to parasites in β -thal trait cells than to those in G6PD⁻ cells, while menadione and high O₂ were inhibitory to both. Thus, riboflavin and menadione have different effects on parasitised G6PD⁻ cells. Both of these agents can participate in single electron redox reactions to generate free radicals^{9,10}, but menadione alone can catalyse H₂O₂ production when reduced by flavin dehydrogenases¹⁰. Indeed, the ability of the malaria parasite to oxidise external NADPH¹¹ suggests that a parasite dehydrogenase may be exposed to the red cell cytoplasm and

Table 2 Multiplication in medium without glutathione

RBC	Multiplication relative to complete medium (mean $\pm$ s.d.)	
	RPMI-1640 -glutathione	High K ⁺ RPMI-1640 -glutathione
Normal	0.84 $\pm$ 0.03 ($n = 4$)	0.70 $\pm$ 0.01 ($n = 2$)
β -thal	0.48 $\pm$ 0.14 ($n = 7$)	0.82 $\pm$ 0.02 ($n = 2$)
G6PD ⁻	0.32 $\pm$ 0.04 ($n = 2$)	
Fetal (HbF)	0.33 $\pm$ 0.18 ($n = 3$)	

Parasites were cultured for 2 d at 17% O₂ in complete RPMI-1640 and then for 2 d in either medium with or without glutathione. In two experiments, growth was also tested in medium without glutathione, but containing high potassium²². All experiments were on samples from different donors.

that menadione contributes to the peroxide load in the red cell by interacting with this enzyme. This interpretation is consistent with the sensitivity of G6PD⁻ cells to H₂O₂ referred to earlier. The sensitivity of thal cells to riboflavin may then be explained by a sensitivity of these cells to free radical mediated oxidation. Vitamin E (α -tocopherol), which prevents the lipid peroxidation catalysed by free radicals¹², was found to completely reverse the sensitivity of thal cells.

Parasitised G6PD⁻ cells were only partially protected by vitamin E, but dithiothreitol (DTT) was more effective. This thiol-reducing agent is expected to reduce glutathione, providing substrate for glutathione peroxidase, and may, like NADPH¹³, protect red cell catalase. Both of these activities would result in lower H₂O₂ levels in the cell and decrease the risk to the G6PD⁻ cell from peroxide production and NADPH utilisation by the parasite. Note that DTT also protected the thal cell from menadione induced damage, presumably by counteracting the H₂O₂ produced by this agent. Mechanisms of oxidant stress and damage are notoriously difficult to identify and prove. The explanation discussed here and shown in Fig. 2 is not meant to be accurate or complete but is offered for discussion and testing.

The situation *in vivo* is not completely comparable to that *in vitro*. Oxygen concentrations of 30% and 25% are not normal physiological conditions. However, riboflavin is found in the circulation and soluble quinones are found in some foods. The observations that these substances potentiate the resistance of variant red cells to malaria infection implicate nutritional interactions in *in vivo* protection. Indeed, such ideas may explain the variable levels of resistance to malaria measured in field studies^{15,16}. In addition, the culture medium RPMI-1640 contains reduced glutathione (GSH). When GSH was left out of the medium, approximating the *in vivo* situation, parasite growth was inhibited in variant cells at 17% O₂ (Table 2) ($P < 0.01$). Growth in HbF-containing fetal cells was also inhibited. HbF

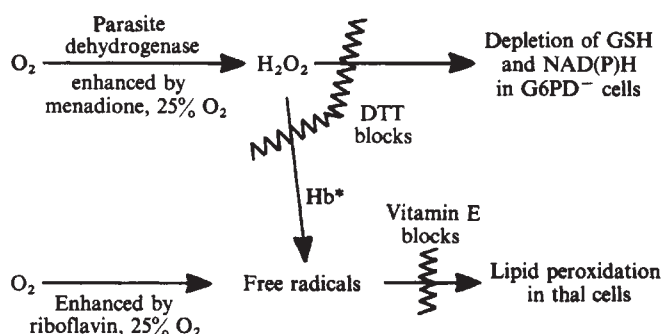


Fig. 2 Hypothetical model for the generation of oxidant damage during malaria infection of variant cells.* The interaction of H₂O₂ and HbF results in free radical production¹⁴. This may be enhanced in thalassaemia or HbF-containing cells.

was initially found to inhibit parasite growth in cultures in medium 199 (0.05 mg per litre GSH)³, but this was not found here in complete RPMI-1640, which contains more glutathione (1 mg l⁻¹). Parasites in HbF cells were only inhibited in the absence of GSH. The sensitivity of fetal cells to lipid peroxidation⁶ is consistent with this observation.

We have now seen that oxidant stress and oxidative damage act to limit malaria parasite multiplication in some red cell variants in conditions similar to those *in vivo*. In G6PD⁻ cells, the production of peroxide by the parasite seems to overload the reducing capacity of the red cell. When catalase (50 µg ml⁻¹) was added to the medium to decrease H₂O₂ levels, parasites in G6PD⁻ cells were protected, whereas those in thal cells were not (data not shown). Although some damage in the G6PD⁻ host is prevented by vitamin E, there are other consequences of H₂O₂ exposure such as the depletion of GSH¹⁷ and NADPH, metabolic requirements of the parasite^{18,11}. *In vivo*, decreased GSH could lead to cross-linking of red cell membrane proteins¹⁹ or Heinz body formation which may result in splenic sequestration and the selective removal from circulation of infected cells. This would explain the paucity of parasitised G6PD⁻ cells seen in the blood of an infected heterozygous mosaic female patient²⁰.

As vitamin E was completely protective in thal red cells, free radical production was thought to lead to lipid peroxidation and membrane damage. To test this hypothesis, parasites in thal trait cells were cultured in a glutathione-free medium containing high (105 mM) K⁺. One of the consequences of oxidative damage to the red cell membrane would have been increased cation permeability²¹. The high K⁺ medium prevents red cell K⁺ loss and maintains intracellular levels even when the membrane leaks. This protected the parasites from inhibition (Table 2). Thus, the primary lesion in the infected thal cell must be to the membrane. The parasite's requirement for a high potassium environment has previously been shown in the sickled HbAS cell which also leaks potassium and inhibits parasite metabolism²².

Although HbF is clearly inhibitory in its own right, the adult thal trait is also resistant. Both the fetal red cell (data not shown) and the thal red cell are sensitive to riboflavin-induced oxidant damage. The basis for this sensitivity is being investigated. Their sensitivity is qualitatively different from that of G6PD⁻ cells; the protection against malaria conferred by these traits may also differ. Nevertheless, the model of oxidant sensitivity presented here provides experimentally verified mechanisms for the resistance of these variants to malaria and for the subsequent enhanced survival of the carriers of these genetic traits.

I thank Drs M. Bestak and H. Chang for thalassaemia trait and fetal cord red cells and the New York Blood Center for samples from blood donors. This work was presented in part at the 1978 meeting of the American Society of Hematology²³.

MILTON J. FRIEDMAN

The Rockefeller University,
New York, New York 10021

Received 10 January; accepted 1 June, 1979.

- Friedman, M. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1994-1997 (1978).
- Pasvol, G., Weatherall, D. J. & Wilson, R. J. M. *Nature* **274**, 701-703 (1978).
- Pasvol, G., Weatherall, D. J. & Wilson, R. J. M. *Nature* **270**, 171-173 (1977).
- Etkin, N. & Eaton, J. in *Erythrocyte Structure and Function* (ed. Brewer, G. J.) 219-232. (Liss, New York, 1975).
- Beutler, E. *J. Lab. clin. Med.* **49**, 84-95 (1957).
- Stocks, J., Offerman, E. L., Modell, C. B. & Dormandy, T. L. *Br. J. Haemat.* **23**, 713-724 (1972).
- Trager, W. & Jensen, J. B. *Science* **193**, 673-675 (1976).
- Motulsky, A. G. & Campbell-Kraut, J. M. in *Proceedings of the Conference on Genetics, Polymorphism, and Geographic Variations in Disease* (ed. Blumberg, B. S.) 159-180 (Grune-Stratton, New York, 1961).
- Massey, V., Palmer, G. & Ballou, D. in *Flavins and Flavoproteins* (ed. Kamin, H.) 349-361 (University Park Press, Baltimore, 1971).
- Misra, H. P. & Fridovich, I. *J. biol. Chem.* **247**, 188-192 (1972).
- Eckman, J. R. & Eaton, J. W. *Nature* **278**, 754-756 (1979).
- Tappel, A. L. *Vitam. Horm.* **20**, 493-510 (1962).
- Eaton, J. W., Boraas, M. & Etkin, N. L. in *Hemoglobin and Red Cell Structure and Function* (ed. Brewer, G. J.) 121-130 (Plenum, New York, 1972).
- Cohen, G. in *Erythrocyte Structure and Function* (ed. Brewer, G. J.) 685-698 (Liss, New York, 1975).

- Livingstone, F. B. A. *Rev. Genet.* **5**, 33-64 (1971).
- Martin, S. K. *et al. Lancet* **i**, 524-526 (1979).
- Cohen, G. & Hochstein, P. *Science* **134**, 1756-1757 (1961).
- Trager, W. *Acta Tropica* **14**, 289-301 (1957).
- Alan, D. W., Johnson, G. J., Cadman, S. & Kaplan, M. E. *J. Lab. clin. Med.* **91**, 321-327 (1978).
- Luzatto, L., Usanga, E. A. & Reddy, S. *Science* **164**, 839-842 (1969).
- Miller, A. & Smith, H. C. *Br. J. Haemat.* **19**, 417-428 (1970).
- Friedman, M. J., Roth, E. F., Nagel, R. L. & Trager, W. *Exp Parasit.* **47**, 73-80 (1979).
- Friedman, M. J. *Blood* **52**, suppl. 1, 96 (1978).

Nucleotide sequences of the attachment sites of bacteriophage Mu DNA

THE temperate bacteriophage Mu has the remarkable ability to insert its DNA in apparently random sites of the *Escherichia coli* chromosome (1-3). All Mu prophages have the same gene order⁴, and the finding that the prophage and the phage maps are identical suggests that specific sites (attachment sites) at the ends of the Mu genome are used for the integration of Mu⁵. Phage particles do not contain the excised, free form of the Mu DNA; instead, the Mu genome is flanked by what seems to be heterogeneous bacterial DNA^{6,7}. The variable DNA at both ends of Mu is thought to be generated when Mu DNA is transposed to many new chromosomal sites during replication and is subsequently packaged into phage particles together with adjacent host DNA. The nucleotide sequence analysis of the ends of Mu DNA reported here substantiates this hypothesis and identifies the attachment sites of Mu. Structural features of the attachment sites are presented.

The material for sequencing was derived in the following way. DNA fragments containing the Mu ends (Mu plus host DNA) were purified from mature phage DNA and inserted into plasmids. The cloning of the *Hind*III fragment, which contains the c-terminal or left end of Mu, has been described elsewhere⁸. Right-end fragments generated by a mixture of *Eco*RI and *Sal*I or *Eco*RI and *Bam*HI were inserted into plasmid pBR322 which had been cleaved with the appropriate restriction nucleases. The structures of these recombinant plasmids are outlined in Fig. 1; plasmids containing the left or right end of Mu are called pLM and pRM, respectively. The sequencing strategy (Fig. 2) involved end-labelling of small juncture fragments with [γ -³²P]ATP and polynucleotide kinase, separation of the two labelled ends by recleavage with restriction nucleases and deduction of the sequence by the chemical modification procedure of Maxam and Gilbert⁹.

The sequence information we have obtained from three plasmids containing the left or right end of Mu demonstrates the validity of the concept that Mu DNA is covalently linked to variable host DNA sequences. For example, sequencing gels of two different left-end juncture fragments from pLM44 and pLM54, both read from the same 5'-labelled *Hpa*II site in Mu DNA into host DNA (Fig. 2), show sequences which match perfectly up to a certain nucleotide but which differ beyond this. In Fig. 3 we have aligned the juncture sequences of pLM and pRM plasmids. For comparison we have included both juncture sequences from a Mu prophage inserted in *lacZ*; details and implications of this latter experiment will be described elsewhere. All pLM plasmids contain a common sequence located next to a variable region different in each plasmid. The same is true for all pRM plasmids. We regard the common sequences as Mu and the variable sequences as host DNA, such that the end points of the Mu genome coincide with the first nucleotides of the common sequence. Although the host DNA sequences are largely different from one another, they show similarities in sequences directly adjacent to the Mu ends. At the left end in two out of four plasmids, the Mu sequences are preceded by

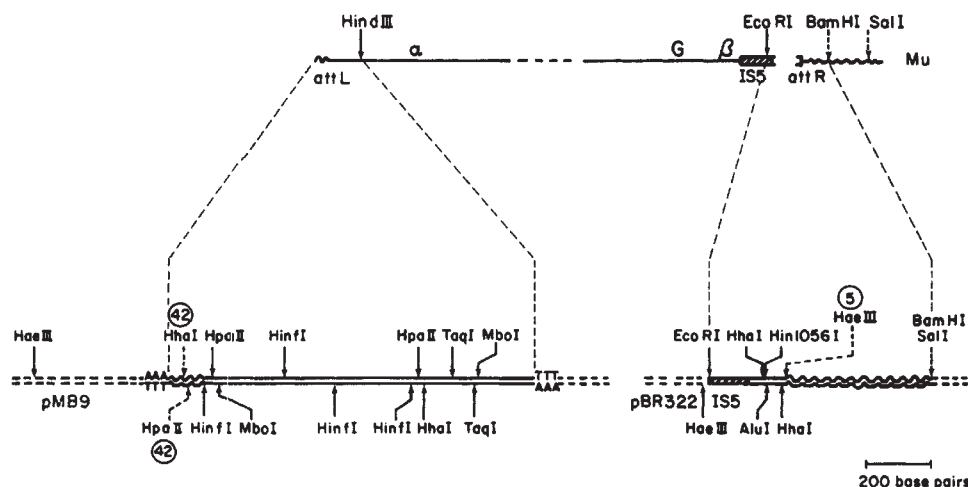


Fig. 1 Cloning of end fragments of the Mu genome and restriction enzyme cleavage map. attL: the leftmost 1,000-base pair long *Hind*III fragment of Mu DNA was isolated and inserted into the *Eco*RI cleavage site of pMB9 after tailing with terminal transferase⁸. A preliminary cleavage map of this fragment has been published previously¹⁹, and we include here only cleavage sites that have been used for the identification of juncture fragments and the sequence analysis. pLM44, pLM54 and pLM42 were used here. attR: Mu 445-8 phage DNA²⁰ carried an additional *Eco*RI cleavage site (R.K. and D.K., unpublished) which is located in the region shown to be IS5 (ref. 21). The rightmost *Eco*RI fragment of Mu 445-8 DNA was isolated, recleaved with either *Sal*I or *Bam*HI at statistically occurring sites in the heterogeneous host DNA and inserted into pBR322 (ref. 22), which had been cleaved with *Eco*RI + *Sal*I

or *Eco*RI + *Bam*HI, respectively. The ligation conditions used were as described by Murray and Murray²³. Three plasmids were analysed, pRM5, containing an *Eco*RI-*Sal*I insert 800 base pairs long, pRM8, containing an *Eco*RI-*Sal*I insert of 2,500 base pairs, and pRM1, containing an *Eco*RI-*Bam*HI insert of 400 base pairs. The heterogeneous *Eco*RI fragment overlapping attR of Mu 445-8 was also cleaved with various restriction enzymes to detect cleavage sites in the constant Mu part, which could then be used to determine the presence of Mu DNA in the pRM chimaeric plasmids. No cleavage sites were detected for *Hpa*II, *Hae*III and *Mbo*I, whereas *Alu*I, *Hha*I and *Hin*1056I cleaved at least once. Solid lines represent Mu DNA, wavy lines indicate host DNA, specific in each plasmid but different from one plasmid to another, and broken lines represent vector DNA. Cleavage sites are indicated by solid line arrows, broken line arrows labelled with numbers indicate that the cleavage site is only present in the plasmid indicated.

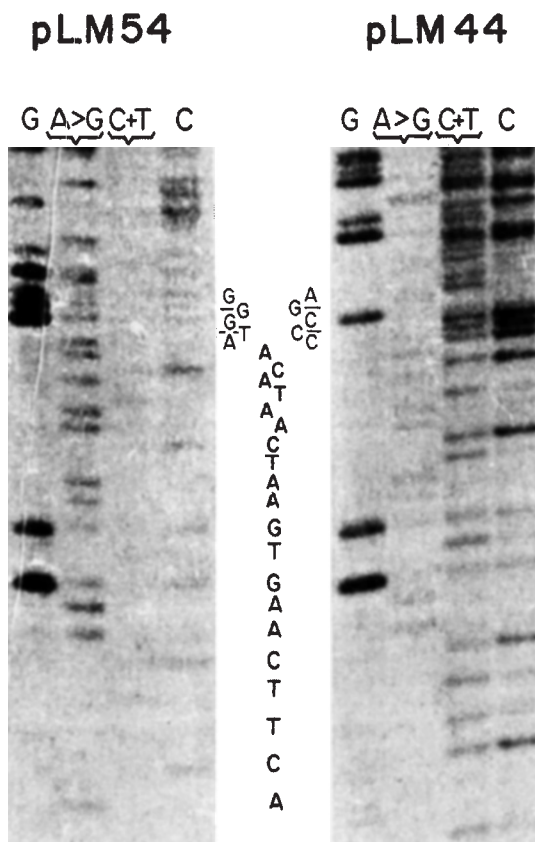
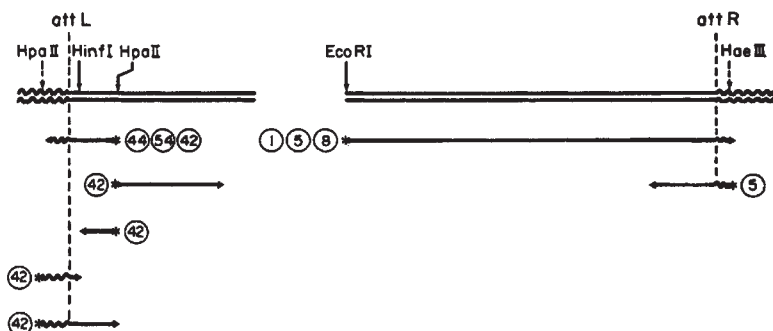


Fig. 2 Sequence analysis. Autoradiographs of sequencing gels showing the attL host DNA-Mu DNA junctures in pLM44 and pLM54. The autoradiographs show parts of the sequences (starting at nucleotide 19 of attL in Fig. 3) read from a labelled 5' *Hpa*II end within Mu DNA through the juncture into the host DNA. Electrophoresis of the DNA fragments generated in the reactions described by Maxam and Gilbert⁹ was on 20% polyacrylamide gels containing 7 M urea. For sequencing the attL juncture fragments of pLM44 and pLM54, the plasmids were digested with *Hpa*II, labelled with [γ -³²P]ATP and polynucleotide kinase⁹, and recut with *Hae*III. The juncture fragments, now labelled only at the *Hpa*II site in the Mu part, were isolated from 2% agarose gels and sequenced. pLM42 contains a *Hpa*II as well as a *Hha*I site in the host DNA. The *Hha*I fragment overlapping the pLM42 attL juncture was isolated after separation on 3.5% polyacrylamide gels; it was then recut with *Hpa*II and end-labelled. The two labelled internal *Hpa*II fragments were isolated from a 5% acrylamide gel, recleaved with *Hin*FI and the products separated on 8% polyacrylamide. All four labelled fragments were sequenced. In addition, the strands of the *Hpa*II juncture fragment of pLM42 were separated as described by Maxam and Gilbert⁹ and sequenced. pRM plasmid DNAs were linearised with *Eco*RI, end-labelled and recleaved with *Hae*III to separate the two labelled ends. The juncture fragments were then purified on 2% agarose gels, re-isolated and sequenced. pRM5 contains an *Hae*III cleavage site in the bacterial DNA very close to the juncture. This site was labelled and the sequence was read over the juncture into Mu DNA. Solid lines indicate Mu, broken lines vector, and wavy lines host DNA, respectively. Asterisks mark the ³²P-labelled ends of a fragment. The arrow indicates the direction and length of the sequence that was read. We consider those sequences which were read from only one strand as preliminary. Numbers in circles refer to plasmid numbers and indicate that fragments from the corresponding plasmid were sequenced.



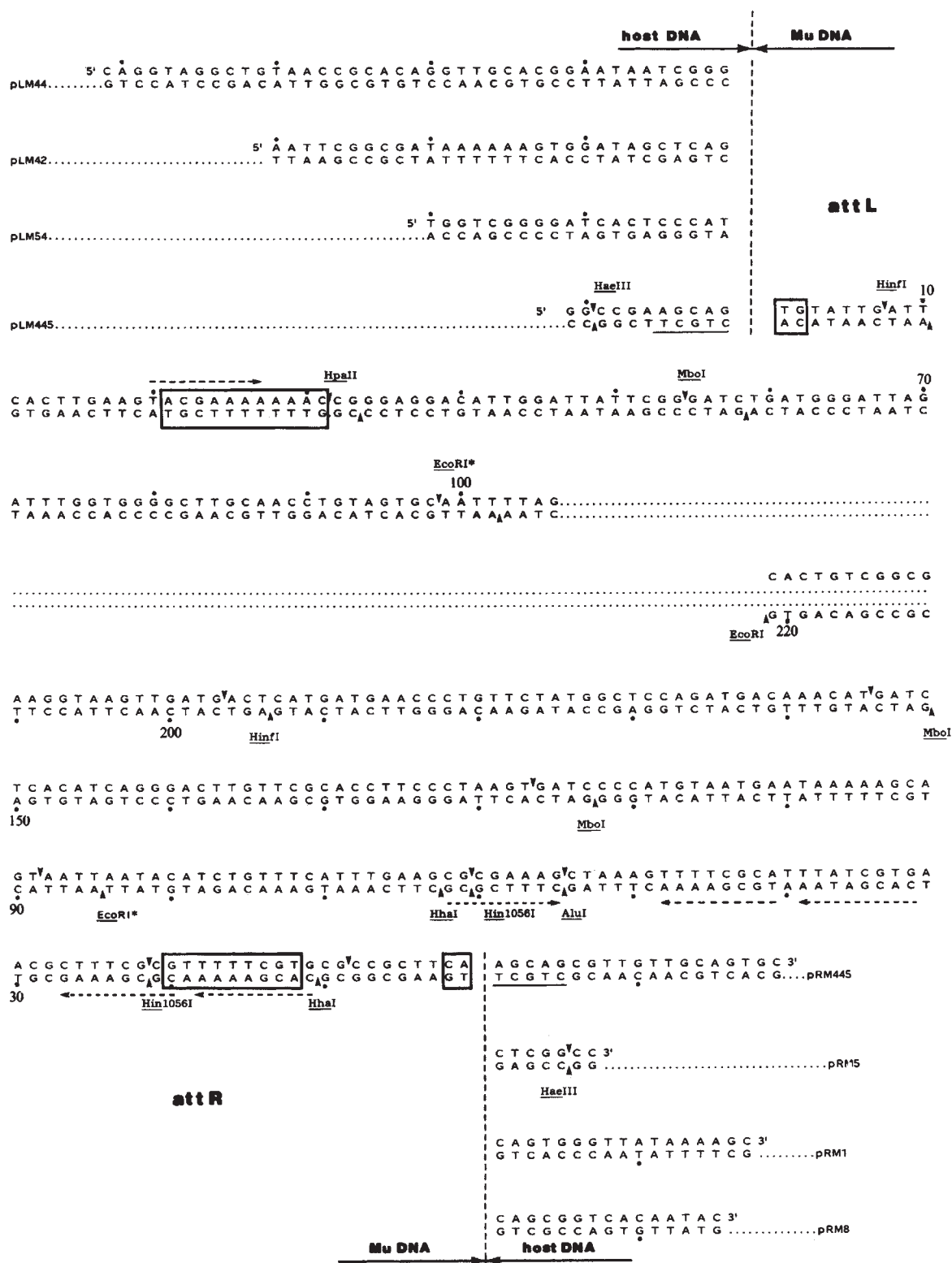


Fig. 3 Nucleotide sequence of the ends of bacteriophage Mu DNA. The central part of the sequence shows the common region of four pLM and four pRM plasmids. The unique sequences for the individual plasmids are shown to the left and right of the common sequence. The junction sequences of pLM445 and pRM445 are included for comparison and will be described in detail elsewhere (manuscript in preparation); they show the junction sequences of a Mu prophage inserted into the *lacZ* gene. Sequences found in inverted order at both ends are in solid line boxes. Broken line arrows indicate the sequence PyPuCGAAAA. The 5-base pair duplication of adjacent host sequences in pLM445 and pRM445 is underlined. The *MboI* and *HinfI* cleavage sites in attR and all *EcoRI** cleavage sites were inferred from the sequence; the presence of all other restriction enzyme cleavage sites have been demonstrated by mapping. Numbers indicate distances from the junctions in base pairs. Note that the attR sequence is derived from Mu 445-8 DNA; Chow *et al.*²⁰ have shown that in Mu 445-8 only ~150 base pairs next to the junction are Mu DNA.

5'CAG and in one case by 5'CAT. At the right end in two of the four junctures the Mu sequence is followed by 5'CAG. We consider this to indicate some specificity of Mu insertion. In fact, it is possible that the specificity in host DNA sequences is more extensive and covers the first nucleotide(s) of the common

sequence so that the definition of the exact end points of the Mu genome becomes somewhat arbitrary. Our positioning of these points, however, has been corroborated by the sequence analysis of two prophage junctures where the host DNA sequence is known (manuscript in preparation).

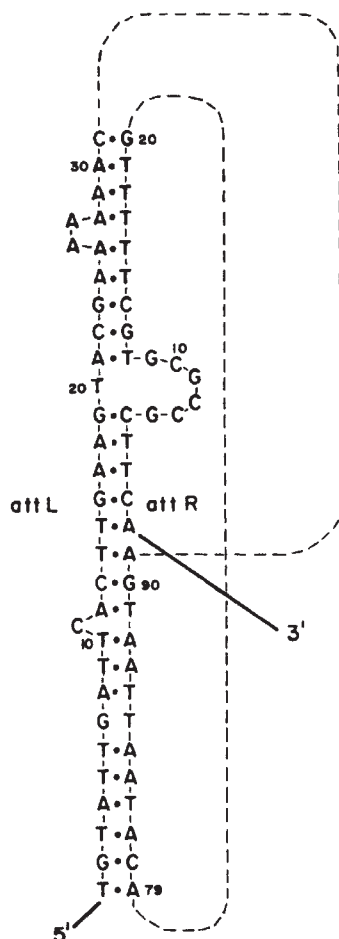


Fig. 4 Similar sequences at Mu attL and Mu attR. To illustrate the arrangement of similar sequences at the two Mu attachment sites, the DNA is presented in a single-strand stem-loop structure. Numbers indicate distances from the Mu-host DNA junctions. Nucleotides 1–20 of attR are paired with nucleotides 15–31 of attL and nucleotides 1–14 of attL are paired with nucleotides 79–91 of attR. Correct base pairing is indicated by dots.

The presence of defined juncture points between Mu and host DNA at both ends of the Mu genome suggests that specific Mu DNA sequences (hereafter referred to as Mu attL and Mu attR for left and right end attachment sites, respectively) at or near the junctures represent recognition sites for enzymes involved in Mu integration. In this respect, it should be noted that only one Mu-specific function, the protein coded for by gene *A*, is implicated in the integration process¹⁰. Assuming that both attachment sites are recognised by this protein we expect to find structural similarities between attL and attR. A first inspection shows that sequence similarities of the two attachment sites are confined to two base pairs at the very end and to a stretch of nine nucleotides (containing one mismatch) located between positions 21 and 31 in attL and between positions 12 and 20 in attR. Moreover, the sequence 5'PyPuCGAAAA which is part of the 9-base pair sequence common to both attachment sites is repeated with minor modifications five times within the first 62 nucleotides in attR but not in attL. This sequence might play a major part in the interaction of protein(s) with the attachment sites. Closer examination reveals that the first 31 nucleotides in attL are almost completely present in attR, but are arranged differently: the first 14 base pairs of attL are found 79 base pairs away from the juncture in attR, whereas nucleotides 15–31 in attL seem to correspond to the first 20 nucleotides of attR (Fig. 4). Whether this sequence arrangement has functional significance or is the result of evolutionary development remains to be determined.

In these structural aspects the Mu attachment sites differ markedly from IS sequences and other transposable elements,

where the ends have been shown to contain more or less identical sequences arranged in inverted orientation^{11–13}. The sequence differences encountered between Mu attL and Mu attR could reflect functional differences of the two attachment sites during integration and in the DNA replication and packaging processes. The insertion mechanisms of Mu and other transposable elements, however, may be closely related, as is indicated by the finding that Mu insertion, like the insertion of IS1 (refs 12, 14), IS2 (ref. 13) and Tn3 (ref. 15) results in a duplication of host DNA at the host target site (Fig. 3)^{16,17}.

The sequences of the ends of Mu DNA shown here and those determined independently by Allet¹⁸ differ substantially. The interpretation of Allet's sequencing data, particularly the location of the Mu ends and the finding of a 5-base pair inverted repeat sequence at the ends, is not consistent with our data. In addition, the sequences beyond the *AluI* cleavage site at the attR end in Allet's sequence are completely different from the sequence we have determined for that region.

This work was supported by NIH grant GM23996 to L. T. Chow, NSF grant GB-32041 and NIH grant 17612 to D. Zipser, NSF grant PCM76-82448 to R. J. Roberts and NIH/NCI Cancer Center Research Grant CA13106 to J. D. Watson.

REGINE KAHMANN

DIETMAR KAMP

Cold Spring Harbor Laboratory,
Cold Spring Harbor, New York 11724

Received 24 April; accepted 18 May 1979.

1. Taylor, A. J. *Proc. natn. Acad. Sci. U.S.A.* **50**, 1043–1051 (1963).
2. Bukhari, A. I. & Zipser, D. *Nature new Biol.* **236**, 240–243 (1972).
3. Daniell, E., Roberts, R. J. & Abelson, J. *J. molec. Biol.* **69**, 1–8 (1972).
4. Abelson, J. *et al. Virology* **54**, 90–92 (1973).
5. Hsu, M.-T. & Davidson, N. *Virology* **58**, 229–239 (1974).
6. Daniell, E., Kohne, D. E. & Abelson, J. *J. Virol.* **15**, 739–743 (1975).
7. Bukhari, A. I., Froshauer, S. & Botchan, M. *Nature* **264**, 580–583 (1976).
8. Zipser, D., Moses, P., Kahmann, R. & Kamp, D. *Gene* **2**, 263–271 (1977).
9. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
10. O'Day, K. J., Schultz, D. W. & Howe, M. M. *Microbiology* (ed. Schlessinger, D.), 48–51 (American Society for Microbiology, 1978).
11. Ohtsubo, H. & Ohtsubo, E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 615–619.
12. Grindley, N. D. F. *Cell* **13**, 419–426 (1978).
13. Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. *Nature* **272**, 414 (1978).
14. Calos, M. P., Johnsrud, L. & Miller, J. H. *Cell* **13**, 411–418 (1978).
15. Ohtsubo, H., Ohmori, H. & Ohtsubo, E. *Cold Spring Harb. Symp. quant. Biol.* **43** (1979).
16. Kamp, D. *et al. Cold Spring Harb. Symp. quant. Biol.* **43** (1979).
17. Allet, B. *Cell* **16**, 123–129 (1979).
18. Allet, B. *Nature* **274**, 553–558 (1978).
19. Kahmann, R., Kamp, D. & Zipser, D. *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) (Cold Spring Harbor, New York, 1977).
20. Chow, L. T., Kahmann, R. & Kamp, D. *J. molec. Biol.* **113**, 591–609 (1977).
21. Chow, L. T. & Broker, T. R. *J. Bact.* **133**, 1427–1436 (1978).
22. Bolivar, R. *et al. Gene* **2**, 95–113 (1977).
23. Murray, N. E. & Murray, K. *Nature* **251**, 476–481 (1974).

Degradation of DL-leucine with longitudinally polarised electrons

MOLECULAR asymmetry represents a striking characteristic of living organisms and their products. The precise mechanism(s) involved in the origin of this asymmetry in nature is not known, although Vester¹ and Ulbricht² proposed that it may have originated through preferential destruction of one enantiomer of a racemic mixture by spin-polarised electrons arising from β -decay, or circularly polarised bremsstrahlung produced during their deceleration. However, tests of this hypothesis have yielded results which were either negative or, at best, of borderline significance³. Recently, however, Bonner *et al.*^{4,5} reported the asymmetric degradation of the amino acid DL-leucine on irradiation with a beam of longitudinally polarised 120-keV electrons (degree of polarisation 10–27%). Electrons polarised

Table 1 Effect of irradiation of samples of DL-leucine with longitudinally polarised electrons with spins antiparallel (AP) and parallel (P) to their momenta*a*, Results of present study

Sample	Degree of electron polarisation (%)	Spin direction	Current density (nA cm ⁻²)	Sample thickness (mg cm ⁻²)	% Decomposition	Enantiomeric* composition (% D-isomer)
1	36	P	690	19	60	49.90 ± 0.16(3)
2	42	AP	555	39	35	50.19 ± 0.15(4)
3	42	AP	915	26	45	50.18 ± 0.28(4)
4	42	P	645	30	45	50.29 ± 0.25(12)
5	42	AP	1,110	31	55	50.20 ± 0.11(8)
6	41	P	795	29	65	50.14 ± 0.24(7)
7	42	P	930	29	75	50.26 ± 0.25(3)
8	42	AP	360	17	65	50.15 ± 0.17(4)
9	45	AP	360	27	40	49.97 ± 0.12(5)
10	45	AP	300	36	40	50.14 ± 0.12(4)
11	47	P	345	29	68	50.24 ± 0.26(9)
12	46	P	390	20	77	49.99 ± 0.30(7)
13	43	P	195	20	75	49.99 ± 0.26(9)
14	43	AP	150	18	70	50.13 ± 0.22(8)
15	43	P	96	23	72	50.14 ± 0.22(4)
16	43	AP	114	23	60	49.92 ± 0.15(8)
17	37	P	93	28	47	49.94 ± 0.16(4)
18	41	AP	111	29	48	49.99 ± 0.10(4)
19	43	P	84	20	65	49.92 ± 0.20(8)
20	43	AP	87	20	78	49.84 ± 0.19(6)
21	42	P	120	26	60	50.08 ± 0.15(8)
22	44	AP	159	18	70	49.96 ± 0.18(4)
23	41	P	27	8.7	65	49.92 ± 0.25(6)
24	43	P	24	14	51	50.10 ± 0.14(8)
25	43	AP	38	8.5	48	49.96 ± 0.11(6)
26	41	P	27	10	60	50.06 ± 0.10(5)
27	40	AP	36	9.2	56	50.08 ± 0.17(5)
Averages: P(14 samples): 50.07 ± 0.03 [†]						AP(13 samples): 50.05 ± 0.03

b, Results of Bonner *et al.*

1	~20	AP	9.4	~20	52.9	49.29 ± 0.20(10)
2	17-23	AP	15.0	~20	50.9	49.57 ± 0.14(18)
3	27	AP	12.5	~6	54.9	49.70 ± 0.16(15)
4	13-17	P	28.8	~20	73.9	50.40 ± 0.12(22)
5	10-17	P	40.0	~20	75.6	50.37 ± 0.14(28)
6	~20	P	14.4	~20	76.3	50.57 ± 0.13(16)
Averages: P(3 samples): 50.45 ± 0.06						AP(3 samples): 49.52 ± 0.12

% Decomposition is estimated from values measured by the enantiomeric marker technique²⁸ on preliminary samples and checked on representative samples (11, 12, 16, 18, 19, 21, 23-27).

* The error is the standard deviation from the mean for the number of analyses given in parentheses.

† The error is $\sigma/\sqrt{n}$ where σ is the standard deviation from the average %D for the *n* P or AP samples.

with their spins antiparallel (AP) to their momenta were reported to cause preferential destruction of the D-enantiomer. Moreover, the L-enantiomer of leucine was reported to be preferentially destroyed by electrons polarised with their spins parallel (P) to their momenta. These results were of special interest as electrons emitted in β -decay have AP spins and also it is the L-isomer of leucine which occurs in nature. The precise physical processes responsible for such an asymmetric degradation remain unclear. We report here that using an electron source⁶ which provides a much higher degree of polarisation (typically 43%), we have reinvestigated the effect of irradiation of DL-leucine with longitudinally polarised electrons. In particular, the increased spin polarisation was expected to result in a more pronounced asymmetric degradation. However, no asymmetric decomposition was observed.

The polarised electron source used has been described elsewhere⁶. Briefly, spin-orientated helium (²S) metastable atoms are produced by optical pumping in a flowing helium afterglow. Electrons generated in chemi-ionisation reactions involving those metastable atoms and a reactant gas are polarised due to spin conservation⁷. The direction of the electron spin polarisation is easily reversed, without changing the beam trajectory, by

reversing the sense of the circularly polarised optical pumping light. The electrons are extracted from the afterglow and formed into a collimated beam by a series of electrostatic lenses. The polarisation of the output beam is determined by measuring the Mott scattering asymmetry when the beam is scattered from a thin gold foil biased at 120 kV (for review see ref. 8). The electrons extracted from the afterglow are polarised transversely as this is convenient for Mott polarisation analysis. A Wein filter⁹ located in the beam line is used to rotate the electron spins to obtain the beam of longitudinally polarised electrons required for sample bombardment.

Solid DL-leucine samples (0.6-2.5 mg) were mounted for irradiation on a thin sheet of aluminium foil backed by a layer of scintillation plastic and biased at 120 kV. Samples were held in place either by an extremely thin collodion film (1-16) or, as used by Bonner *et al.*^{4,5}, in a collodion matrix (17-27). Samples were irradiated with the beam polarised with electron spins parallel and antiparallel to their momentum vectors for target thickness ranging from 8.5 to 40 mg cm⁻². Combinations of electron beam current and irradiation times were varied to obtain sample decomposition (as measured by the enantiomeric marker technique¹⁰) of 35-80%. Although the electron source

used can produce microampere beam currents, we kept currents low to prevent heating of the sample. Despite the wide range of current densities ($24\text{--}1,110\text{ nA cm}^{-2}$) no significant variation was observed in the decomposition rate ($0.016\text{ mg nA}^{-1}\text{ h}^{-1}$), which compared favourably with that of Bonner *et al.*^{4,5} ($0.015\text{ mg nA}^{-1}\text{ h}^{-1}$).

After irradiation, samples were dissolved in 3 M HCl in 2-propanol (2 ml) and maintained at 100°C for 3 h. After solvent evaporation, the resulting 2-propyl esters were treated overnight at room temperature with a 1:1 mixture of trifluoroacetic anhydride and methylene chloride (2 ml). The enantiomeric composition of the resulting *N*-trifluoroacetyl 2-propyl esters was determined by gas-liquid chromatography (GLC)¹¹ on a column ($2.4\text{ m} \times 2\text{ mm}$) of SP-300 (5% *N*-lauroyl-L-valyl-*tert*-butylamide on Supelcoport; 100–120 mesh) using a Hewlett-Packard model 402 unit coupled to a Hewlett-Packard model 3370B digital electronic integrator. The precision and sensitivity of the method were confirmed using leucine of known enantiomeric composition processed in a manner identical to the irradiated samples. With 49.5% D-enantiomer our technique measured $49.56 \pm 0.1\%$ (5 analyses); with 50.5% D-enantiomer it measured $50.54 \pm 0.08\%$ (7 analyses).

Each irradiated sample of DL-leucine provided sufficient material for several determinations of enantiomeric composition. Before each determination either a sample irradiated with unpolarised electrons or a non-irradiated sample was analysed as a control to identify instrumental difficulties. To preclude bias GLC was carried out with the sense of polarisation of the bombarding electrons concealed.

As Table 1 shows, in spite of the wide range of experimental parameters investigated and attempts (samples 23–27) to duplicate as closely as possible the conditions reported by Bonner *et al.*⁴, we observed no asymmetric decomposition of DL-leucine by longitudinally polarised electrons. Indeed we observed no single case of preferential decomposition of one of the enantiomers of DL-leucine on irradiation with polarised electrons.

We thank A. R. Harrison for technical assistance. This investigation was supported in part by the NIH (HL-15376), the Department of Energy (Office of Basic Energy Sciences) and the Robert A. Welch Foundation (C-583 and C-704).

L. A. HODGE
F. B. DUNNING
G. K. WALTERS

Physics Department,
Rice University

R. H. WHITE
G. J. SCHROEPFER JR

Biochemistry Department,
Rice University,
Houston, Texas 77001

Received 16 August 1978; accepted 22 May 1979.

1. Vester, F. seminar, Yale Univ. 7 Feb. (1957).
2. Ulbricht, T. L. V. *Q. Rev.* **13**, 48–60 (1959).
3. Ulbricht, T. L. V. *Nature* **258**, 383–384 (1975).
4. Bonner, W. A., Van Dort, M. A. & Yearian, M. R. *Nature* **258**, 419–421 (1975).
5. Bonner, W. A., Van Dort, M. A., Yearian, M. R., Zeman, H. D. & Li, G. C. *Israel J. Chem.* **15**, 89–95 (1977).
6. Hodge, L. A., Dunning, F. B. & Walters, G. K. *Rev. Sci. Instrum.* **50**, 1 (1979).
7. Kehrer, P. J., Dunning, F. B., O'Neill, M. R., Rundel, R. D. & Walters, G. K. *Phys. Rev. A* **11**, 1271–1278 (1975).
8. Kessler, J. in *Polarized Electrons* (Springer, Berlin, 1976).
9. Boersch, H., Geiger, J. & Stickel, W. *Z. Phys.* **180**, 415–424 (1964).
10. Bonner, W. A. *J. chromat. Sci.* **11**, 101–102 (1973).
11. Bonner, W. A., Van Dort, M. & Flores, J. J. *Analyt. Chem.* **46**, 2104–2107 (1974).

REPLY—We are indebted to Drs Walters and Hodge for making available to us an earlier version of the above manuscript in which only the first 22 experiments were reported. At our suggestion that these experiments were not comparable to

ours due to the much higher current densities used, they undertook experiments 23–27, where the experimental parameters were apparently much closer to those we used. It is therefore particularly puzzling to us that these last five experiments also failed to show any asymmetric degradation.

In our linear accelerator irradiations we also occasionally (in about one-third of our experiments) observed no asymmetric degradation whatsoever. When an asymmetric effect was observed, however, it was consistently in one direction for AP electrons (three experiments; 43 GLC analyses) and in the opposite direction for P electron (three experiments; 66 GLC analyses), with statistical significance (*t* test) at the 99.9+ % confidence level. Clearly, there are unknown operational parameters over which we, at least, had incomplete control in our experiments.

Since the initial publication of our positive results we have reported^{1,2} the discovery of the racemisation of amino acids by ionising radiation. It seems likely that such radioracemisation must also compete with asymmetric degradation² in irradiation experiments such as ours and those of Hodge *et al.*, whether because of the electrons themselves or their bremsstrahlung. Could it be that unknown experimental parameters might tip the balance between a slight asymmetric bias in such irradiations, or (with overriding radioracemisation) none at all? Future experiments may shed light on this or other possible reasons for the unexplained discrepancies in our results and those of Hodge *et al.*

W. A. BONNER

M. R. YEARIAN

M. A. VAN DORT

Stanford University,
Stanford, California 94305

1. Bonner, W. A. & Lemmon, R. M. *J. molec. Evol.* **11**, 95 (1978).
2. Bonner, W. A. & Lemmon, R. M. *Bioorg. Chem.* **7**, 175 (1978).

Errata

In the letter 'Calcium conductance of acetylcholine-induced endplate channels' by P. D. Bregestovski, R. Miledi and I. Parker, *Nature* **279**, 638–639, the greek letter tau (τ) was omitted from six places in the text. These are: para 2 line 13; para 6 lines 1, 4, and 6 (twice); para 9 line 4. These errors are corrected in reprints of the article.

In the article 'Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex' by C. D. Gilbert and T. N. Wiesel, *Nature* **280**, 120–125, in para 4 line 8, for diaminobutyric acid read 3,3'-diaminobenzidine.

Corrigendum

Since publication of the letter 'A Moravian age for the "younger Moines" of central and western Scotland' by M. A. J. Piasecki and O. van Breemen, *Nature* **278**, 734–736, the term 'Grampian group' has been changed to 'Grampian Division'. This follows discussions with an IGS working party and others. The term 'Division' is as used by Johnstone *et al. Mem. Am. Ass. Petrol. Geol.* **12**, 159–180 (1969).

matters arising

Mechanisms for radiation-induced shrinkage of voids

EVANS¹ has recently reported on the radiation-induced shrinkage of voids in molybdenum and its alloy TZM (Mo, 0.5% Ti, 0.1% Zr). Electron microscope measurements made on 3-mm disks irradiated in a fast reactor at $450 \pm 25^\circ\text{C}$ reveal that initially produced swelling of 0.57 and 0.34% at a total neutron dose of $3.5 \times 10^{26} \text{ m}^{-2}$ equivalent to 17.5 displacements per atom (d.p.a.) reduced on re-irradiation to 0.20 and 0.01% at a neutron dose of $8 \times 10^{26} \text{ m}^{-2}$ or 40 d.p.a. in molybdenum and TZM respectively. A related though opposite behaviour is found in the vacancy type loop structure, typically 5 nm in diameter, whose concentration increased to $6 \times 10^{22} \text{ m}^{-3}$ and $1.3 \times 10^{23} \text{ m}^{-3}$ at the highest dose level in molybdenum and TZM respectively.

Three mechanisms can give rise to negative swelling rates. First, solute atom segregation at the void surface considered by Marwick² which makes the voids biased against vacancies, a possibility discussed by Evans. Second, shrinkage due to thermal emission from voids having a radius less than a critical value³⁻⁶, a mechanism which requires high temperatures and is unlikely to be important in the present context. The third mechanism^{5,6}, which is applicable at relatively low temperatures of 450°C and in conditions where vacancy loops are produced, seems to be pertinent for the observations reported for molybdenum and TZM. The trapping of a fraction ϵ of the radiation-produced vacancies as vacancy loops or as impurities would require that a certain minimum dislocation density ρ_{crit} should develop so that the bias Z_i would allow a sufficient number of vacancies to go to neutral sinks like voids. An expression for ρ_{crit} involving the recombination parameter α , dose rate K and vacancy and interstitial diffusion constants D_v and D_i respectively can be obtained within the framework of the BEK model⁷. The voids will therefore shrink if the total dislocation density due to network and loop dislocation is less than^{5,6}

$$\rho_{\text{crit}} = (\alpha K / D_v D_i)^{1/2} \epsilon / (Z_i - 1)$$

The high vacancy migration energy of about 2.0 eV for molybdenum will make

D_v small and ρ_{crit} large at low temperatures. Though accurate values of the parameters are not known an estimate using $\alpha/D_i = 10^{20} \text{ m}^{-2}$, $K = 10^{-6} \text{ d.p.a. s}^{-1}$ and $\epsilon \sim (Z_i - 1) \approx 0.05$ would give ρ_{crit} in the range $10^{15} - 10^{16} \text{ m}^{-2}$. This compares very favourably with the observed vacancy loop densities of $9 \times 10^{14} \text{ m}^{-2}$ and $2 \times 10^{15} \text{ m}^{-2}$ for molybdenum and TZM respectively calculated from the loop size and concentration values reported. The mechanism therefore seems to be relevant for explaining the radiation-induced shrinkage of voids observed by Evans. The reason that the voids should initially grow at these temperatures and dislocation density values needs careful examination and explanation.

K. KRISHAN

R. V. NANDEKAR

Reactor Research Centre,
Kalpakkam 603 102,
Tamil Nadu, India

1. Evans, J. H. *Nature* **278**, 728-729 (1979).
2. Marwick, A. D. *J. Phys.* **F8**, 1849-1861 (1978).
3. Krishan, K. *Rep. RRC-12* (Reactor Research Centre, Kalpakkam, 1976).
4. Hayns, M. R. *AERE Rep. R-8806* (1977).
5. Krishan, K. & Nandedkar, R. V. *Proc. Indo-German Seminar on Radiation Damage*, 156-199 (Reactor Research Centre, Kalpakkam, 1977).
6. Krishan, K. & Nandedkar, R. V. *Pramana* (in the press).
7. Bullough, R., Eyre, B. L. & Krishan, K. *Proc. R. Soc. A* **346**, 81-102 (1975).
8. Brimhall, J. L., Simonen, E. P. & Kissinger, H. E. *J. nuclear Mater.* **48**, 339-350 (1973).

EVANS REPLIES—Krishan and Nandedkar correctly point out that negative swelling rates can arise in a situation where a fraction of the radiation-produced vacancies are effectively removed from the system in the form of vacancy loops. Their expression showing the importance of dislocation density is for the mutual recombination dominant regime. For situations where recombination can be ignored but where thermal emission of vacancies from the loops is still negligible it is easy to show that the condition for void shrinkage is given by $\epsilon > (Z_i - 1) / (Z_i + y)^{-1}$, where Z_i is the bias and y is the ratio of void sink strength to dislocation sink strength. This expression may be more appropriate to the molybdenum case if a vacancy migration energy of 1.3 eV (ref. 1) is accepted.

To apply the mechanism suggested by Krishan and Nandedkar to my void shrinkage results² one must, as they point out, explain the growth of voids in the earlier stages of irradiation. In theory this is not too difficult: if one considers the

expression above, an initial high dislocation density (and low y) could allow growth of voids. If the dislocation density remains constant the associated increase in y actually leads to a novel saturation condition for swelling but a subsequent drop in dislocation density would lead to the required void shrinkage. However, this picture is too inconsistent with experimental observations to be seriously considered for the molybdenum situation. One example from the void shrinkage results is the simultaneous large increase in dislocation density contrary to the requirements above, but more generally there is the problem of void growth in molybdenum at temperatures down to 330°C (ref. 3) ($\sim 0.21 T_m$) even when the void sink strength is dominant ($y \sim 10$), and where the Krishan and Nandedkar assumption $\epsilon(Z_i - 1)^{-1} \sim 1$ would imply negative swelling rates. The only easy way out of this dilemma is to accept that in molybdenum $\epsilon(Z_i - 1)^{-1} \ll 0.1$ and in contrast to materials such as 316 stainless steel^{4,5}, assume that the vacancy loops play a negligible part in swelling processes below the $0.3 T_m$ threshold. The differences between cascade collapse behaviour in molybdenum and f.c.c. metals during ion bombardment⁶ apparently support this viewpoint.

The suggestions of Krishan and Nandedkar may have applications elsewhere, for example, in a thermal cycling situation ((A. J. E. Foreman, personal communication) but cannot easily explain the molybdenum and TZM results. I still feel that the explanation already provided^{3,7}—where the segregation of a transmutation impurity (probably technetium) to void surfaces imposes an effective bias on the voids simply by having a low exchange probability for a vacancy—fits the experimental facts reasonably well. Note that this explanation does not depend on the reverse Kirkendall flux that Marwick⁸ has suggested for concentrated alloys. However, it does closely fit the surface energy barrier model recently suggested by Fikus and Johnson⁹.

J. H. EVANS

Metallurgy Division,
AERE Harwell,
Oxon, UK

1. Balluffi, R. W. *J. nuclear Mater.* **69**, 240-263 (1978).
2. Evans, J. H. *Nature* **278**, 728-729 (1979).
3. Bentley, J. thesis, Univ. Birmingham (1974).
4. Bullough, R., Eyre, B. L. & Krishan, K. *Proc. R. Soc. A* **346**, 81-102 (1975).

5. Foreman, A. J. E. & Makin, M. J. *J. nuclear Mater.* **79**, 43–57 (1979).
6. English, C. A., Eyre, B. L., Summers, J. & Wadley, H. *Proc. Int. Conf. on Fundamental Aspects of Radiation Damage in Metals* 751006-P2, 910–917 (ERDA Conf., 1975).
7. Evans, J. H. *Int. Conf. on Irradiation Behaviour of Metallic Materials for Fast Reactor Core Components*, Corsica 1979 (in the press).
8. Marwick, A. D. *J. Phys.* **F8**, 1849–1861 (1978).
9. Fikus, J. H. & Johnson, R. A. *Radiat. Effects* **40**, 63–70 (1979).

Possible Sun-weather correlation

AS I have recently published¹ a re-examination of some of Xanthakis' work², in which he claimed remarkably high correlations between solar activity and an index of zonal precipitation (for example, $r = +0.77$ in the zone 70°–80° N and $r = -0.71$ for 60°–70° N), I feel I should comment on the recent exchange between Xanthakis and Gerety^{3,4}.

My re-examination was prompted by the unusually high correlations claimed, the questionable discussion by Xanthakis of 'X-wise distributions'⁵ and, more particularly, by the arbitrariness of the derivation of zonal mean $R - R_0$ values. Xanthakis defined $R - R_0$ for each year and station as the actual total precipitation, R , at that station in that year minus the lowest annual precipitation, R_0 , reported for that same station over all the years considered. In the calculation of the zonal means Xanthakis averages together stations having very different mean $R - R_0$ values, with no attempt at normalisation (apart from assigning almost arbitrary R_0 values). Furthermore the stations come in or drop out in random fashion according to whether records are available. Thus, even if a perfect correlation between $R - R_0$ values and solar activity existed at each station (which is certainly not the case), the averaging process would be expected to inject discontinuities whenever a station record began or ended. In this respect Xanthakis' method should lead to poor correlations. This makes the claimed high correlations even more remarkable, and indeed puzzling.

The data for 70°–80° N were re-evaluated because that zone contains the least data, Xanthakis published a full tabulation of his data for that zone², and his result for that zone was an apparently highly significant correlation ($r = +0.77$) which did not change sign. Like Gerety *et al.*⁶ I found that Xanthakis had used only part of the available data, and my analysis using additional data does not support the claimed highly significant correlation. I obtain a correlation of +0.38 which, considering that the data have been smoothed, is barely significant at the 95% confidence level. Xanthakis rightly points out that a different data sample will lead to a different result. However, the great reduction in the observed correlation using the larger sample certainly throws into question the significance of Xanthakis' result.

More importantly, unlike Gerety *et al.*, I chose only to use data as described by Xanthakis, who stated (or certainly implied) that he used all the available data published in *World Weather Records*⁷ except that "only stations with more or less extensive series of precipitation data have been considered"². An examination of his Table 2a shows that he used stations with as little as 6 years (Ingoy) or 8 years (Myggbukta) of data.

Examination of *World Weather Records* reveals a complete set of data for Isfjord Radio, Spitsbergen for 1912–24, which Xanthakis omitted even though this meant that for 1912–21 his 'zonal mean' was based on only one station. For the interval 1951–60 there are data for five additional stations—Isachsen (10 yr), Sachs Harbour (5 yr), Clyde (9 yr), Barter Island (10 yr), Chokurdakli (2 yr), and Vize Island (8 yr). All but Chokurdakli lie between the meridians 156° W through 0° to 80° E as stipulated by Xanthakis.

The issue is thus not whether different samples of stations were used, but rather why Xanthakis used the particular stations (which give such remarkable results) and did not use others which detract from the significance of the results. Furthermore, why was no mention made and no reason given for omission from the zonal mean of apparently good data (available from the same data source) from long-record stations such as Isfjord Radio, Isachsen and Barter Island, when shorter records were used from Ingoy and Myggbukta? Why does the choice of stations for inclusion in the zonal means consistently lead to higher correlations than those found using larger data samples?

In an attempt to answer these questions I have tested for the possibility that the correlation coefficients found by Xanthakis using his samples of stations and those found by using the larger available data samples are drawn from the same population by a random or chance selection of stations. According to Snedecor⁸ this is done by applying a Z transform to the correlation coefficients and a Student t test to the difference between the Z values.

Table 1 summarises the results for three latitude belts which have been indepen-

dently assessed by Pittock¹ and Gerety *et al.*⁶, following the procedures outlined by Xanthakis². Assuming N independent data pairs as indicated for each latitude belt, the probability P that each pair of samples of stations is drawn at random from the same population is shown. N has been reduced to allow for the computed autocorrelations in the data (including smoothing) according to the standard formula given by Quenouille⁹. Taken separately, the result for each of the three latitude zones suggests that Xanthakis' selection of stations (a selection which gives apparently significant correlations) was probably not by chance. Combining the results, it seems that the joint probability that Xanthakis' favourable selection was by chance is less than 1 in 30,000.

I am grateful to Dr E. J. Gerety for supplying the results of his computations.

A. BARRIE PITTOCK

Laboratory of Tree-Ring Research,
University of Arizona,
Tucson, Arizona 85721

and
CSIRO Division of Atmospheric Physics,
Mordialloc 3195, Australia

1. Pittock, A. B. *Rev. Geophys. Space Phys.* **16**, 400–420 (1978).
2. Xanthakis, J. *Solar Activity and Related Interplanetary and Terrestrial Phenomena* (ed. Xanthakis, J.) 20–47 (Springer, New York, 1973).
3. Xanthakis, J. *Nature* **275**, 775 (1978).
4. Gerety, E. J. *Nature* **275**, 775 (1978).
5. Roberts, W. O. *WMO Bull.* **23**, 205–206 (1974).
6. Gerety, E. J., Olson, R. H. & Roberts, W. O. *Nature* **272**, 231–232 (1978).
7. *World Weather Records*, Smithsonian Miscellaneous Collections, Vols **79**, **90**, **105**; U.S. Department of Commerce, Vols 1–6.
8. Snedecor, G. W. *Statistical Methods* 5th edn, 178 (Iowa State College Press, 1956).
9. Quenouille, M. H. *Associated Measurements*, 168 (Butterworths, London, 1952).

XANTHAKIS REPLIES—Pittock claims that the X-wise distribution of the reported correlations between $R - R_0$ and I_a (ref. 1) is questionable. The same dispute was put forward previously by Roberts². However, in the light of recent analysis by Gerety *et al.*³ it seems that in the zone 50°–60° N their correlation between $R - R_0$ and I_a has changed sign from negative (1885–1913) to positive (1914–60)³. This change of sign of the correlation, which is

Table 1 Correlation coefficients, r , between smoothed $R - R_0$ and Xanthakis' solar activity index I_a , for three latitude zones over the years indicated, as obtained by Xanthakis² with his selection of stations, and those obtained by Pittock¹ (for 70°–80° N) and Gerety *et al.*⁶ (for 60°–70° N and 50°–60° N) using larger available data samples

Zone	Years	r			
		Xanthakis	Larger samples	N	P
70°–80° N	1912–60	+0.77	+0.38	19	~0.07
60°–70° N	1882–1960	–0.71	–0.35	31	~0.05
50°–60° N	1914–65	+0.79	+0.20	21	~0.009

P is the probability that the data samples in each latitude zone are drawn at random from the same population, where N independent data pairs are assumed to exist. N has been reduced to allow for the computed autocorrelations in the data (including smoothing) according to the standard formula given by Quenouille⁹.

also observed at other latitude zones as well⁴, is what we have originally called X-wise distribution¹.

The values of R_0 , at individual stations, were not assigned "almost arbitrarily" but, as Pittock also mentioned, they refer to the lowest annual precipitation reported for the same station over all the years considered.

Pittock re-examined the results for the latitude belt 70°–80° N which, as he agrees, contain the least data. But this zone was only examined¹ to get a general idea of the behaviour of the quantity $R - R_0$ over the polar regions.

As regards the stations Barter Island (70°08' N) and Clyde (70°27' N), which are located at the boundary of the latitude belts 70°–80° N and 60°–70° N, these were grouped with the data of the 60°–70° N belt (see Table 3 of ref. 1) and not in the 70°–80° N, because their $R - R_0$ values are negatively correlated with I_a index, that is, as is the case for the corresponding data of the stations in the belt 60°–70° N. It is obvious that if in the very few stations (1–6) of the zone 70°–80° N, two more stations are added, negatively correlated with I_a , the correlation coefficient will be reduced significantly. Also, as King points out, "Analyses combining data from zones exhibiting opposite solar-cycle effects will inevitably lead to the erroneous conclusion that no solar-cycle effect exists, as will an analysis of rainfall data from regions situated between zones in which the sunspot effect is opposite"⁶.

The omission of the stations Isachsen (10 yr), Sachs Harbour (5 yr) and Vize Island (8 yr) during the interval 1951–1960 was considered necessary for homogeneity purposes of the record to avoid an abrupt increase of the number of stations in that period.

With regard to the results of Table 1 in Pittock's analysis, I should note that no selection of stations was made for the zone 50°–60° N, for which all stations published in *World Weather Records* were included in the calculations.

Only in the belt 60°–70° N, 20 Siberian stations (long 65°–170° E) were studied separately, because of the different behaviour of their $R - R_0$ values with respect to the corresponding values of the remaining 20 stations in that zone, as can also be seen in Fig. 5 of ref. 1. This separation in my opinion was necessary.

Finally these research studies have been extended to include all the northern and southern latitude zones more comprehensively^{4,7,8}.

PITTOCK ADDS—Xanthakis' X-wise distribution is questionable because it involves a post-hoc decision for each data point as to whether a positive or negative correlation is appropriate. This adds an extra degree for each data point and enables any random scatter of data to be turned into an impressive apparent correlation. Results obtained by such a process have no statistical significance.

Xanthakis defined R_0 as the lowest annual precipitation reported at that station over all the years considered, but in fact he rounded off R_0 values to the nearest 100 mm below the lowest annual precipitation for most of the stations, but sometimes to the nearest 100 mm above (for example, Myggbukta and Vardo) and in one case to 300 mm at Jan Mayen where the lowest annual precipitation was 197 mm in 1926. "Almost arbitrary" seems to be a fair description of this process.

Xanthakis now admits that the stations included in his zonal means were selected on a post-hoc basis according to the signs of their individual correlations of $R - R_0$ with I_a . This clearly invalidates Xanthakis's claimed levels of statistical significance. As indicated in my discussion of X-wise distributions, any set of random data can be classified into sub-groups on a post-hoc basis so as to give apparently significant correlations within sub-groups.

Xanthakis has still not explained why he omitted from his calculations for the zone 70°–80° N the data for Isfjord Radio, Spitsbergen for 1912–24 even though he used data from the same station for 1925–60.

When someone comes up with an *a priori* basis on theoretical grounds for grouping stations or data by area, zone or era, and then shows statistically significant correlations within such previously defined sub-groups, we may get somewhere. Selection of data on an arbitrary post-hoc basis proves nothing.

JOHN XANTHAKIS

Research Center for Astronomy
and Applied Mathematics,
Academy of Athens,
Athens 136, Greece

1. Xanthakis, J. *Proc. 1st Eur. Astr. Meet.*, 1 (Springer, Berlin, 1973).
2. Roberts, WMO Bull. 23, 205–206 (1974).
3. Gerety, E. J., Olson, R. H. & Roberts, W. D. *Nature* 272, 231–232 (1978).
4. Xanthakis, J. *Solar Activity and a Global Survey of Precipitation* (Research Center for Astronomy and Applied Mathematics, Academy of Athens, 1975).
5. Xanthakis, J. & Gerety, E. J. *Nature* 275, 775 (1978).
6. King, J. W. *Aeronaut. Astronaut.* 13, No. 1 (1975).
7. Xanthakis, J. & Tritakis, B. *Praktika de l'Academie d'Athènes* Vol. 51, 600, Contr. No. 47 (1977).
8. Zerefos, C. & Tritakis, B. *Praktika de l'Academie d'Athènes* (in the press).

Use of the solar limb effect to test photon decay and cosmological redshift theories

CRAWFORD¹ suggested photon decay as the result of a gradual energy loss of photons due to their interaction with a curved space-time caused by tidal stresses on the finite spread of the photons. He attempts to explain with his theory the Hubble redshift, the solar limb effect, and the photon decay observed from the Pioneer-6 spacecraft when close to the Sun. The main support for his theory comes from the apparent excellent agreement of the prediction of his theory and the observed

solar limb effect. The solar limb effect is the systematic shift of the wavelengths of solar Fraunhofer lines towards red when going from the centre to the limb of the Sun after solar rotation effects have been removed. The solar limb effect has been used in many other more or less exotic theories involving photon decay, cosmological redshifts and so on^{2–12}. It is therefore of interest to summarise the properties of the limb effect and to compare them, in this instance, with the Crawford theory.

Any theory put forward to explain the limb effect should explain the following properties. (1) After removal of the Einstein gravitational redshift corresponding to 636 m s⁻¹, solar spectral lines show a wavelength shift of Δ to the blue at the disk centre. (2) The magnitude of Δ depends strongly on the strength of the line used^{13,14} on the position in the line at which the shift is measured¹⁵ and on the excitation and ionisation potential of the line^{16,17}. For weak lines Δ corresponds to ~400 m s⁻¹; for strong lines to ~0 m s⁻¹. (3) When going from centre to limb the wavelength stays nearly constant until one approaches the limb closely. At 0.2 solar radius from the limb the wavelength starts shifting towards the red. The absolute shift reaches 0 m s⁻¹ at ~0.02 solar radius from the limb and shows a slight redshift (~0.1 Δ) closer to the limb¹¹, which probably disappears at the limb itself. This redshift has been called 'super-gravitational redshift'. (4) In sunspots the limb effect is absent ($\Delta = 0$)¹⁸.

It was recently shown¹⁹ that all these properties can be explained qualitatively in terms of so-called convective blue shifts resulting from the bright upward-moving convective elements in the solar atmosphere (granules) contributing more to the average radiation than the dark downward-moving elements. Quantitatively, the calculated convective shifts fall short by a factor 2–3 of the observed ones, but that discrepancy may be caused by an inadequate granulation model based on only partially resolved observations of the solar granulation.

Crawford's theory predicts a redshifting of lines when going from the centre of the Sun to the solar limb corresponding to 911 m s⁻¹, which is independent of line strength and position in the line. The same shift should be present in sunspots. Crawford tests his theory using data obtained for relatively weak lines by Plaskett²⁰, Adam²¹ and Higgs²² after correction of the observed wavelengths for a postulated average radial flow of matter into the Sun of 360 m s⁻¹. He obtains good agreement between the variation of wavelengths between centre and limb predicted by his theory and the 'corrected' data. Had he used the limb effect in strong lines or in sunspots the agreement would have vanished. Had he not postulated the 360 m s⁻¹ inflow, a procedure which in our view has no justification, the agreement

would also have disappeared. Crawford also predicts a redshift of the solar lines at the solar limb of 911 m s^{-1} in addition to the 636 m s^{-1} Einstein gravitational redshift (1.547 km s^{-1} in total). But observations at the solar limb give a value close (within $\sim 50 \text{ m s}^{-1}$) to the gravitational redshift leaving an order of magnitude disagreement with the Crawford theory.

There is in our view no justification of Crawford's theory in terms of the observations of the solar limb effect. A major part, and possibly all, of the solar limb effect can be explained by simple physical processes in the solar atmosphere involving convective, wave and pressure shifts without taking recourse to any of the many exotic theories proposed. Observational and theoretical studies of the solar limb effect, from the point of view of solar atmospheric dynamics are obviously most important in setting limits to the magnitude of the effect predicted by theories of photon decay.

The Sacramento Peak Observatory is operated by the Association of Universities for Research in Astronomy, under contract AST 78-17292 with the NSF.

JACQUES M. BECKERS
LAWRENCE E. CRAM

Sacramento Peak Observatory,
Sunspot, New Mexico 88349

1. Crawford, D. F. *Nature* **277**, 633-635 (1979).
2. Freundlich, E. F. *Nachrichten Akad. Wiss. Göttingen* **95**-102 (1953).
3. Freundlich, E. F. *Proc. phys. Soc. A* **67**, 192-193 (1954).
4. Freundlich, E. F. *Phil. Mag.* **45**, 303-319 (1954).
5. Gillieson, A. H. *Can. J. Phys.* **43**, 57-73 (1965).
6. Proisy, P. *C. r. heb. Séanc. Acad. Sci., Paris* **B265**, 1145-1148 (1967).
7. Ferencz, Cs. & Tarcsei, Gy. *Planet. Space Sci.* **19**, 659-667 (1971).
8. Gasanalizade, A. G. *Solar Phys.* **20**, 507-512 (1971).
9. Gasanalizade, A. G. *Soln. Dann.* No. 12, 101-103 (1972).
10. Kipper, A. *Publ. Tartu Astrofiz. Obs.* **41**, 20-29 (1973).
11. Pecker, J. C., Tait, W. & Vigier, J. P. *Nature* **241**, 338-340 (1973).
12. Smith-Hänni, A. L. *Helv. phys. Acta* **48**, 548-551 (1975).
13. O'Brien, J. E. *Solar Phys.* **19**, 314-322 (1971).
14. Beckers, J. M. & de Vegvar, P. *Solar Phys.* **58**, 7-15 (1978).
15. Adam, M. G., Ibbetson, P. A. & Pettford, A. D. *Mon. Not. R. Astr. Soc.* **177**, 687-708 (1976).
16. Salman-Zade, R. K. *Sov. Astr.* **13**, 466-472 (1969).
17. Lambert, D. L. & Mallia, E. A. *Solar Phys.* **3**, 499-504 (1968).
18. Beckers, J. M. *Astrophys. J.* **213**, 900-905 (1977).
19. Beckers, J. M. & Nelson, G. D. *Solar Phys.* **58**, 243-261 (1978).
20. Plaskett, H. H. *Mon. Not. R. Astr. Soc.* **179**, 339-350 (1977).
21. Adam, M. G. *Mon. Not. R. Astr. Soc.* **119**, 460-474 (1959).
22. Higgs, L. A. *Mon. Not. R. Astr. Soc.* **121**, 421-435 (1960).

The problem of thrown string

OUR attention has been drawn to the problem of the thrown string^{1,2} by Bass and Bracken³, who related the probability density of the distance between the ends of the string when in space to the volume of the ellipsoid in which the string must be contained. We⁴ have already suggested a modification to this theory in an attempt to explain the experimental results of Synge². However, we have recently reconsidered the problem.

We felt that in order to test the validity of the different possible models, an experimental distribution was required

Table 1 Frequency distributions for $\lambda = D/L$

Interval	Frequency 'springy' string	Frequency flexible string
0.0-0.1	344	379
0.1-0.2	641	763
0.2-0.3	739	884
0.3-0.4	662	654
0.4-0.5	485	428
0.5-0.6	481	284
0.6-0.7	391	199
0.7-0.8	427	131
0.8-0.9	450	119
0.9-1.0	400	49
Total	5,020	3,890

based on a large set of observations. The problem of generating the necessary data was solved by posing the problem of the thrown string to our first year engineering students as an exercise in data collection, descriptive statistics and fitting models to data. Each student had to obtain observations for at least a hundred throws of the string. The students were divided into two groups, one used a rather 'springy' red plastic, waxed, string, the other a flexible white cotton string of the same physical dimensions. All the strings were 54 cm long. The pooled data for each group is given separately in Table 1 and the corresponding histograms are shown in Fig. 1.

The density function proposed by Bass and Bracken is shown in Fig. 1. This shows that the model does not give a realistic description of the physical situation.

Observing the thrown string lying on the table it seemed that the steady motion of a point moving along the string from one end towards the other would be some sort of continuous two-dimensional random walk. The problem is then not unlike that studied by Broadbent and Kendall⁵ concerning the wanderings of

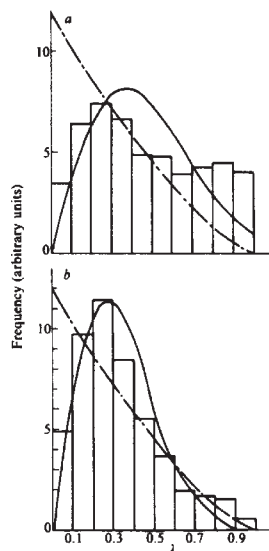


Fig. 1 Histogram and density functions for the 'springy' string (a); and for the flexible string (b). Bass and Bracken model (dashed line); Random walk model (solid line).

the larvae of the helminth *Trichostrongylus retortaeformis*. They assumed that the orthogonal X and Y components of position were independent random functions of time with zero mean and equal variance. X and Y were then assumed to be Wiener processes. This was not reasonable because it implies that the velocities in the x and y directions would be white noise. Perhaps it would have been more reasonable to have assumed that X and Y were the outputs of identical third order lagging systems excited by separate independent white noises. However, provided the systems were linear, this would not affect the result that after a given time X and Y would have independent normal distributions with zero mean and equal variance.

By analogy, our new hypothesis for the thrown string is that the distance D between the two ends will have probability density

$$f(D) = \frac{D}{\omega L^2} \exp\left(-\frac{D^2}{2\omega L^2}\right)$$

where L is the length of the string and ω is a constant dependent on the nature of the string and the experimental conditions. A similar result was reported by Kingman⁶ using a slightly different model.

This density function is compared with the experimental results in Fig. 1 which shows that the proposed model reflects the general nature of the results reasonably well. The discrepancies between the theory and the experiment seem to be due to the lack of ability of the model to describe the upper tail of the distribution. We feel that the large frequencies at the upper tail are primarily a consequence of the impact of the string with the table resulting in a straightening out of the string; the effect is most noticeable in the 'springy' string.

We conclude first that the Bass and Bracken³ and our previous⁴ theories are false in that they are not borne out by experimental results. Second, a good approximation to the distribution is apparently given by the random walk model described above provided the string is not springy. The degree of flexibility of the string may be allowed for by suitable choice of the one parameter in this distribution. More experiments with different non-springy flexible strings are required to verify this second conclusion.

We thank the staff and students who helped with the experiment.

A. P. ROBERTS
D. SPREVAK

Department of Engineering Mathematics,
The Queen's University of Belfast,
Ashby Institute, Belfast,
Northern Ireland, UK

1. Synge, J. L. *Math. Gaz.* **52**, 165 (1968).
2. Synge, J. L. *Math. Gaz.* **54**, 250-260 (1970).
3. Bass, L. & Bracken, A. J. *Nature* **275**, 205-206 (1978).
4. Sprevak, D. & Roberts, A. P. *Nature* **277**, 153-154 (1979).
5. Broadbent, R. & Kendall, D. G. *Biometrics* **9**, 460-466 (1953).
6. Kingman, J. F. C. *Adv. Appl. Prob.* **9**, 431 (1977).

reviews

A physicist who enjoys it

Rudolph Peierls

What Little I Remember. By Otto R. Frisch. Pp.219. (Cambridge University Press: Cambridge, 1979.) £9.50

MOST of nuclear physics has today become "big science", requiring large and expensive accelerators, with complex detection equipment and computers, used by large teams of research workers. Its senior practitioners have to spend much of their time and their energy on administration and organisation. This image of the new scientist does not fit Otto Robert Frisch at all. His very readable reminiscences are those of an individualist, at his best when working and thinking on his own, or with perhaps one close collaborator.

This characteristic comes out very clearly, together with the basic happiness of a man who enjoys what he is doing. Scientists tend to be driven by a variety of motives including, to varying degrees, ambition, a sense of duty to the community, curiosity about nature and its laws, and enjoyment in the design and execution of experiments. All these no doubt have some reality for Frisch, but the pleasure of doing what he does is clearly dominant. Ambition is the least of his concerns.

All this goes for his many other interests. Besides being an outstanding experimenter, with a flair for the simple and decisive experiment, he has many other interests and talents. This breadth probably owes much to his background. He grew up in Vienna, in a family belonging to the cultural tradition of early twentieth-century Vienna. His aunt, the famous nuclear physicist Lise Meitner, was his mentor in much besides physics. Next to physics his strongest line is music. He is a pianist of near-professional standard, and his hearing has absolute pitch. But here again he is less concerned with demonstrating his brilliance at the piano than in the pleasure of making music, and he can enjoy playing with others who may be much below his level in skill, as long as they share his love for music and his pleasure in making it.

He also likes drawing, and after every meeting the papers in front of him are covered with portraits or caricatures. Some of these are reproduced in the book. They are all recognisable likenesses, and some bring out the personality of the subject excellently. His drawing is not of the standard of his music or his physics, but his enjoyment in the activity is again evident.

Presumably the drawings reproduced in the book are the most successful ones of

many. The written characterisations of people, which, in his preface, he compares with the pencil sketches, also vary very much in depth. Niels Bohr and Lise Meitner come very much alive, as does George Placzek, with whom Frisch worked much in Copenhagen. But very little emerges of Chadwick, with whom he was closely associated in Liverpool and later in Los Alamos, beyond an account of their first encounter; and of some other colourful personalities there are only pale shadows.

"non-Aryan", and there followed the typical odyssey of a displaced academic of the time. A year at Birkbeck College in London under Blackett gave him the first acquaintance with life in England. This was the time when artificial radioactivity was discovered, and Frisch made his first venture into nuclear physics: he invented a clever device to observe very short-lived activities, and was able to discover new species of radioactive nuclei. He stayed in London only a year, and was then invited to Bohr's Institute in Copenhagen, where



George Placzek (1905-1955)

Sketch by O. R. Frisch

As a physicist, Frisch studied in the University of Vienna, where he received his PhD (he does not mention the subject of his thesis). After a year in a small firm making scientific instruments he was offered a position in the Physikalisch-Technische Reichsanstalt in Berlin, and after three years moved to Hamburg to work under Otto Stern, the authority on atomic beam research. That much migration was normal at a time when modern physicists formed a close-knit but widely-spread family.

But in 1933, when Hitler came to power, he had to be dismissed from his post as a

he remained until 1939.

This was a period of great excitement in nuclear physics. Fermi had started using the newly discovered neutron as a probe for studying nuclei, and neutron physics became one of the major subjects in many laboratories, including Copenhagen. Frisch became involved in this work, originally because he had a well-running counter suitable for detecting the activities produced by neutron bombardment. He describes the way the work was done, including the preparation of neutron sources by mixing radium and beryllium. A

particularly exciting moment was Niels Bohr's first pronouncement of the idea of the compound nucleus, as a comment after a seminar talk which highlighted the difficulties of the then conventional view in accounting for neutron resonances.

Then, in 1938, came the discovery by Hahn and Strassmann that radioactive barium was one of the products of the bombardment of uranium with neutrons. Frisch was visiting Lise Meitner in Sweden when she received the news of Hahn's result, and there is a fascinating account of the discussion which finally led them to realise that the uranium nucleus was undergoing "fission" (a term chosen by Frisch to describe the phenomenon). Before long he was back in Copenhagen and was doing an experiment which verified the suggested explanation.

By this time War was imminent, and Denmark was not a very safe place for a refugee from Hitler, so Frisch moved to Birmingham. He started some experiments relating to the fission problem which was still one of his main interest, but in a department not well equipped for nuclear physics, and with most of its members working on radar, progress was slow. He continued thinking about the implications and possibilities, including nuclear energy. Bohr had shown that a violent explosion was not possible with natural uranium, but Frisch started to question whether the



Frisch with Lise Meitner

possibility of separating the uranium isotopes in substantial quantities was really ruled out. He and Peierls concluded that a nuclear weapon did not seem to be impossible, and wrote a memorandum which led to an atomic weapons project being started in Britain. From then on most of his effort was devoted to this project. It was an odd situation that he and several others working on one of the most secret military projects were technically enemy aliens, and he tells of a number of ludicrous consequences of this. It soon emerged that he could work more usefully in Liverpool, where Chadwick's laboratory had a small cyclotron and was not preoccupied with other War research. He worked there, through the period of heavy air raids, until 1943, when it was decided to discontinue the British project during War time and send to the United States all those who could be useful to the American atom bomb work. So Frisch went to Los Alamos, and he gives a vivid description of life and work in this odd secret town. Very sensibly he was not attached to one of the large research groups there, but had a roving commission, helping with experimental troubles wherever appropriate.

One of his other accomplishments at Los Alamos was to learn driving, and he tells the story of his first accident, which brought him a broken rib and other minor injuries. He does not seem to remember another accident which became legendary at Los Alamos. Pulling aside to clear another car on a narrow lane without footpath, he knocked over an Army girl. The girl spent a few days in hospital and complained that Frisch had never visited her or shown interest. This was reported to him, and he promptly appeared in hospital, apologising and explaining that he had been concerned. "I enquired of the doctor

about about your health and was glad to hear that there was nothing seriously wrong with you, mainly the effect of the fright"

After the War he became head of the Nuclear Physics Division of AERE, Harwell, the new atomic energy laboratory. According to his account of this period, he left the running of the division to his deputy, who enjoyed administration, and continued his individual research. Perhaps his diffident description exaggerates his lack of application to the work of running the division, but it is in character that he was not at his best in this kind of responsibility.

He had been in Harwell less than two years when he was offered the Jacksonian Professorship in Cambridge. The book is exceedingly brief about the 25 years he spent in Cambridge before retiring. He explains in the preface that his memory, like that of many others, is more complete about earlier events than about the more recent ones, and that it is embarrassing to write about people with whom one is still in touch. So the whole period, which includes his marriage and the growth of his two children, gets only a few pages. The story ends with a description of another ingenious device, a machine for measuring and evaluating bubble-chamber tracks, which he designed; he is now a partner in the firm manufacturing this commercially.

Three chapters of the book, *Atoms, Nuclei and Research Resumed*, contain expositions of parts of modern physics by way of background to the accounts of his own involvement with it. Being a physicist myself, I am not qualified to judge the appropriateness of these pieces. They are written with the clarity and simplicity to be expected from the author of *Meet the Atoms*, yet I wonder whether the reader who does not know the background can

Blackie

The Structure of Matter

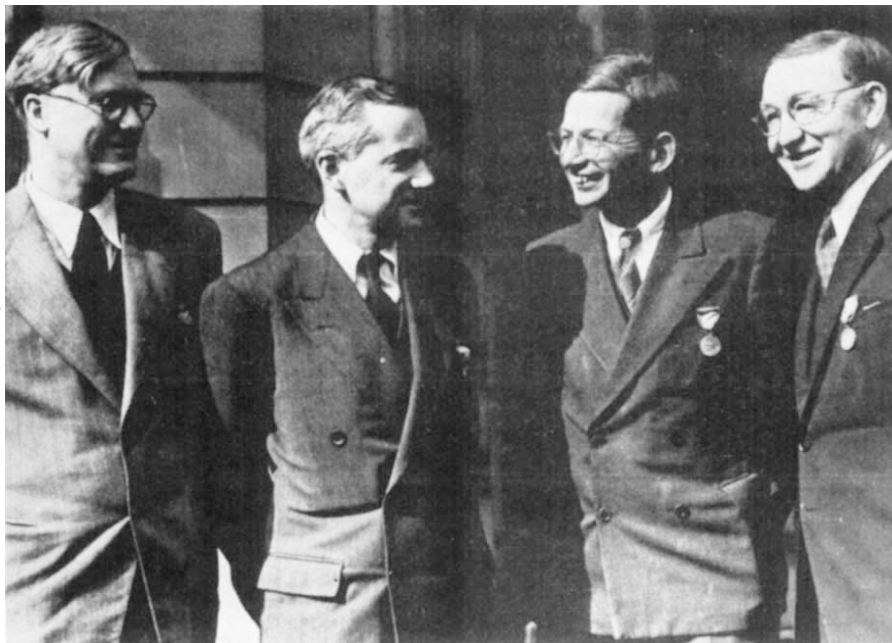
Robert M. Turnbull

A concise introduction to atomic, nuclear and particle physics for students in the early part of an undergraduate course. Deals with the principal topics in atomic and nuclear physics and goes on to describe the most recent findings in particle physics. Includes a problems section with answers.

256 pp 88 illustrations
ISBN 0 216 90753 5
July 1979 £7.95 net



Blackie & Son Ltd.
Bishopbriggs,
Glasgow G64 2NZ, U.K.



Posing for the Press, about 1946

L to R: William Penny, Frisch, Rudolf Peierls, John Cockcroft

really assimilate enough information to help him with the biographical chapters which are the real meat of the book, and which seem quite intelligible on their own.

This is a happy book, from which the author's personality and his enjoyment of physics, of music, of life, emerges clearly. It is also a portrait of the pre-War world of physics, of the days of small numbers and

small apparatus, of times when a physicist could think of an ingenious experiment today and set it up tomorrow. ☐

Sir Rudolf Peierls was Professor of Theoretical Physics at the University of Oxford until his retirement in 1974, and was previously Professor of Mathematical Physics at the University of Birmingham from 1937-63.

Gelada baboons

Ecological and Sociological Studies of Gelada Baboons. Edited by M. Kawai. Pp. 344. (S. Karger: Basel, Munchen, Paris and London, 1978). DM148; \$74.

In a variety of ways, gelada baboons are one of the most unprimate-like primates: they are almost exclusively terrestrial; they feed extensively on grass; and they aggregate in large herds of unstable membership. Consequently, they are a species which might be expected to shed extensive light on relationships between ecological variables and behaviour and on the functional significance of behavioural differences between primates and ungulates.

The volume recently edited by Kawai provides a broad description of the ranging behaviour, social organisation and feeding ecology of the species in the Gich area of the Simien National Park in Ethiopia, but does little to provide satisfactory explanations of the geladas' many surprising characteristics. Four observers habituated several herds and learnt to recognise all individuals in one of them. As a result, they were able to examine both detailed social relationships among individual members of the harem-groups which aggregate to form herds, and relationships between neighbouring herds.

Twelve chapters, by combinations of the

four authors, cover herd dynamics, inter-harem relationships, social behaviour within harems, the development of social behaviour and group structure, reproductive behaviour, communication, ranging behaviour and feeding ecology. Particularly in areas where it concentrates on answering questions rather than broad, quantitative description, the book reveals fascinating insights into the behaviour of the species: its description of herd stability, of grooming relationships within harems and of food intake in relation to abundance, provide novel coverage of important topics.

Sadly, the reader's interest is titillated more often than it is satisfied. Why do females periodically discipline their harem males by chasing them up cliffs? Why do harem masters tolerate the presence of extra males in their harems which, even if they are temporarily inhibited from breeding, are likely to contest their hegemony in future? And why do male geladas not control their females in as chauvinistic a fashion as male hamadryas baboons — which share a similar environment and social system? Attention to specific questions of this kind would have provided the book with a stronger connecting theme and reduced the extent to which it overlaps the previous volume on geladas in the same series.

T. H. Clutton-Brock

T. H. Clutton-Brock is at King's College Research Centre, Cambridge, UK.

Published today:

Food, Energy and Society

David Pimentel and Marcia Pimentel

Resource and Environmental Sciences series

This book examines the relationship between food, energy and society. The authors analyse and compare the ways in which different societies, both 'developed' and 'undeveloped', use different kinds of energy in producing their food. They also investigate the effects of substituting other resources (such as labour or land) for some of this energy, and assess the nutritional quality of a variety of diets.

The aim of this book is to explore the interdependencies of food, energy and their impacts on society. These analyses, it is hoped, will be a basis for planning and implementing policies of individuals and nations as they face the inevitable dilemma of how everyone can be fed, given the limited resources of the earth.

Paper £3.95 Illustrated



Edward Arnold

41 Bedford Square
London WC1B 3DQ

Circle No. 28 on Reader Enquiry Card.

Bring Your Library Up-To-Date With This ANS Edition...

J.M. Cleveland
MONOGRAPH

The Chemistry Of Plutonium

Approx. 680 pages 6" x 9"
Hardbound \$49.00

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Please take my order for _____ book(s).

Payment in full enclosed.

Bill me and I accept postage and handling charges.

Name _____

Organization _____

Street _____

City _____

State _____ Zip _____

Circle No. 29 on Reader Enquiry Card.

©Macmillan Journals Ltd 1979

Trace elements in metabolism

Trace Elements and Iron in Human Metabolism. By A.S. Prasad. Pp. 380. (Plenum: New York, 1979). Available in UK, Middle East and Africa from Wiley (Chichester, UK) at £17.50; \$35.

If a monograph is defined as a systematic and exhaustive examination of a particular subject, then this book, contrary to the statement in the Foreword, cannot be so described. It contains information about not only the well-established (chromium, copper, fluorine, iodine, manganese and selenium) and newer (nickel, vanadium, silicon and tin) trace elements, but also magnesium, iron and certain toxic metals. This last section is highly selective in its choice of metals (cadmium, lead and mercury) and seems quite out of place in the present volume; the brief chapters are superficial and contain statements that are either incorrect, or likely to mislead. Thus, for example: "hypertensive subjects have abnormally high concentrations of renal cadmium" (p354); "A similar protective effect of excess zinc against lead toxicity was noted in young horses" (p365); and, "The physiocochemical classification of mercury does not correlate well with its toxic properties" (p371).

Even the review of the field of essential metals is an enormous task for a single author, and although Dr Prasad has produced an informative summary, inevitably many of the recent developments have been omitted. The chapter on chromium (confined essentially to the glucose-tolerance factor), selenium (devoted largely to glutathione peroxidase) and the newer trace elements, for example, contain only 2, 2 and 1 reference(s), respectively, to work published later than 1975, whilst

in the chapter on iron (a detailed, 60-page, review) 94% of the 428 references are to publications before 1971. Also the references to published work are perhaps inevitably for such a large subject, rather selective. Nevertheless, the book provides a clear account, not by any means exhaustive, of certain aspects of the action of trace elements aiming to link molecular biology, biological and animal experimentation with human nutrition and clinical experience. This is a highly commendable aim and the book can be recommended to provide such a broad view which, if depth is required, will require supplementation with further reading. The individual chapters on copper, iron and zinc, in particular, are extremely thorough and informative.

The very readable text is, in some cases, rather repetitious largely due to the subdivisions in each chapter, such that the experimental work in animals is often separated from observations in man. Furthermore, the considerations of the functions and mechanism of action of the

trace elements are somehow diluted by sections on toxicity and therapeutic measures. The book is well produced, but for such a wide general coverage of the subject, a more detailed index might have been an advantage.

With the provisos mentioned above, this book can be recommended for those who wish to quickly obtain a general view of the biology of trace elements with particular reference to man. Contrary to the title, the volume is not restricted to human metabolism, but at least in certain sections also deals extensively with studies in experimental and domestic animals. Thus, it covers much but not all, of the ground already surveyed by Underwood in the fourth edition (1977) of his book *Trace Elements in Human and Animal Nutrition* (Academic; New York, San Francisco and London).

N. Aldridge
M. Webb

N. Aldridge and M. Webb are at the MRC Toxicology Unit, Carshalton, UK.

Electromagnetic wave propagation

Methods in Electromagnetic Wave Propagation. By D.S. Jones. Pp. 887. (Clarendon/Oxford University Press: Oxford, 1979.) £22.50.

THE field of electromagnetism has, as the author claims, undergone considerable development in recent years. Motivated by such fields as satellite communications, radar and geophysical prospecting the search for better methods of solution to important problems in the areas of electromagnetic wave propagation and scattering has intensified. Existing classical methods of applied analysis have been extended, complemented or indeed supplanted by new numero/theoretic methods based upon modern applicable mathematics. Professor Jones, himself a leading researcher in the field, is to be congratulated on presenting these methods in a systematic manner and critically reviewing their application.

The book is very large, containing over 800 pages in 9 chapters. The first chapter is introductory and contains the mathematical features used in the remainder of the book. The problems of propagation in waveguides are discussed in chapters 2, 3 and 4. Solutions are found by finite differences (chapter 2) and by variational methods (chapter 4). As these problems can often be reduced to finding eigenvalues of operators, this is described in some detail in chapter 3. The implementation of variational methods to provide numerical bounds on approximate solutions is the main concern of chapter 5.

The advent of the digital computer has transformed the antenna design field and classically 'difficult' problems are now amenable to numerical analysis. Both wire and solid antennas are treated in chapter 6 where integral equations of various types are studied from analytical and numerical points of view. Guidance on methods for dealing with transient phenomena is given in chapter 7.

The most exciting material in the book for the reviewer is in chapter 8 where the geometric theory of diffraction (GTD) is developed from first principles. Very few authoritative accounts of this subject are available and the author has devoted one quarter of his book to it. The use of GTD in the design of reflector antennas has become increasingly important in recent years. Finally the problems of inverse scattering, that is, describing the source knowing its scattered field, is discussed in chapter 9. Nearly 400 references are listed, although few appear for 1977 and none afterwards.

It is perhaps churlish to mention omissions in a book of this length but no space has been found for recent work on general reflector synthesis involving Monge-Ampere equations or the use of complex coordinates and stereographic projection in geometrical optics.

The quality of the material and the clarity of its exposition makes this an important reference book for research workers and lecturers (there are numerous exercises) in the electromagnetic field.

B.S. Westcott

B.S. Westcott is Senior Lecturer in Applied Mathematics in the Faculty of Mathematical Studies, University of Southampton, UK.

● *The Sense of Order: A Study in the Psychology of Decorative Art* (Phaidon: Oxford, 1979), by E.H. Gombrich (for review, see *Nature*, 279, 653; 1979), has been published in the United States by Cornell University Press (Ithaca, New York) at \$38.50.

● *The Menace of Atomic Energy* (W.W. Norton: New York, 1977), by Ralph Nader and John Abotts (for review, see *Nature*, 274, 95; 1978) has now been published in the UK by Melbourne House (London) at £7.95.

● *Rutherford and Physics at the Turn of the Century*, edited by M. Bunge and W.R. Shea (for review, see *Nature*, 279, 741; 1979), was published by Neale Watson Academic (New York) and Dawson (Folkestone, UK), and not as stated.